# National Assessment Governing Board

**Ad Hoc Committee on Measures of Postsecondary Preparedness**

**Thursday, November 15, 2018**

**2:00 – 4:00 pm**

**Finalizing Recommendations**

## Agenda

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>2:00 – 2:10 pm</td>
<td>Welcome and Overview of Committee’s Charge</td>
<td><em>Terry Mazany, Committee Chair</em></td>
</tr>
<tr>
<td>2:10 – 2:30 pm</td>
<td>Reflections on the Young Adult Expert Panel Meeting</td>
<td><em>(summary provided separately)</em></td>
</tr>
<tr>
<td>2:30 – 3:45 pm</td>
<td>Discussion Draft Recommendations Report &amp; Operational Implications</td>
<td><em>(draft report provided separately; draft appendix attached)</em></td>
</tr>
<tr>
<td>3:45 – 4:00 pm</td>
<td><strong>ACTION:</strong> Committee Decision to Finalize Recommendations and Report them to the Board</td>
<td><em>Terry Mazany</em></td>
</tr>
</tbody>
</table>
## Table of Appendices

- Appendix A. Charge to the Committee ................................................................. A-2
- Appendix B. Exploratory Approach Overview .................................................. B-4
- Appendix C. Summary Answers for Research Questions ............................... C-6
- Appendix D. Expert Panel: Industry ................................................................. D-15
- Appendix E. Expert Panel: Higher Education .................................................. E-26
- Appendix F. Expert Panel: Futurists ................................................................. F-39
- Appendix G. Expert Panel: State Assessment Directors ............................... G-85
- Appendix H. Expert Panel: Youth ................................................................. H-93
- Appendix I. Literature Review: Work of the Future ...................................... I-95
- Appendix J. Literature Review: Skills of the Future ....................................... J-114
- Appendix K. Literature Review: State Indicators of College and Career Preparedness ...... K-142
Resolution: The Executive Committee’s Charge to the Ad Hoc Committee on Measures of Postsecondary Preparedness

Whereas, on November 18, 2016, the National Assessment Governing Board unanimously approved the Strategic Vision to guide its work through the year 2020; and

Whereas, the Strategic Vision established a Board priority (SV#10) to “Develop new approaches to measure the complex skills required for transition to postsecondary education and career”; and

Whereas, on August 3, 2017, the Governing Board Chair created the Ad Hoc Committee on Measures of Postsecondary Preparedness to pursue this priority; and

Whereas, the Governing Board Chair tasked the Executive Committee to establish the charge to guide the Ad Hoc Committee on Measures of Postsecondary Preparedness;

Therefore, the Executive Committee resolves that:

1. The Ad Hoc Committee on Measures of Postsecondary Preparedness shall review existing research, collect expert testimony, and prepare recommendations for the Governing Board’s consideration to achieve Strategic Vision priority #10.

2. While the current legislation guiding the National Assessment of Educational Progress (P.L. 107-279) should provide parameters for the approaches to accomplish this priority, the Ad Hoc Committee on Measures of Postsecondary Preparedness may consider options that could require amendments to current legislation.

3. The Ad Hoc Committee on Measures of Postsecondary Preparedness will report its recommendations to the Governing Board no later than the November 2018 Board meeting.
Appendix B. Exploratory Approach Overview
To address its charge, the National Assessment Governing Board’s Ad Hoc Committee on Measures of Postsecondary Preparedness considered the trends that most likely will shape the future, and thereby shape the skills and knowledge that students will need to develop. Through meetings with expert panels and commissioning focused research papers, conducted with the support of its contractor HumRRO, the Committee pursued the answers to the following three research questions:

1. **Work of the future (readiness for what?):** What are we, as a nation, preparing students for? Changes in the workplace are not only inevitable, but are accelerating, driven by technological advances, demographic shifts, and social changes. The growing prevalence of self-driving vehicles, the more widespread use of robots, and advances in artificial intelligence are signs of existing innovations poised to dramatically change the jobs available to young Americans. Young Americans hold different expectations about work, and the ways in which people connect and communicate with each other are all changing. How will the workplace change given these emerging technologies? How will our communities change given these trends?

2. **Requisite skills for future work (skills for what?):** With a better understanding of the future workplace, we can better understand the skills that young Americans will need to succeed. But should we consider more than just workplace skills? What about skills like citizenship and financial literacy? How do these skills factor into the question of measuring postsecondary preparedness?

3. **Measures of preparedness (measures for what?):** Finally, what metrics exist to capture the skills that young Americans will need in the workplace, for their roles in their communities, and in their personal lives? Can such metrics include data from sources in addition to or instead of assessments? Additionally, what metrics do not exist but are needed to help the nation better understand if students are prepared as they exit high school, regardless of which paths they take—through college or other postsecondary learning experiences or directly to the workforce?

The following pages summarize the key findings from the various research and expert consultations; more detailed summaries are provided in the subsequent appendices.
**Work of the Future – Readiness for What? (Q1)**

Students beginning their K-12 education in 2017 are on course to graduate in 2030. Upon graduating high school, these students will face many choices for pursuing work and further learning. Traditionally, high school graduates have chosen between enrolling in postsecondary education or directly entering the workforce. Of 2017 high school graduates, for example, 67% enrolled in a postsecondary institution, 22% entered the work force, and 11% did neither.\(^1\) Postsecondary education has typically been a four-year college, two-year college, or technical school. Entry-level jobs are those that permit the employee to enter the workforce with little experience or education, often with the goal of moving onto a career path through valuable on-the-job experience. The military is an alternate pathway, providing training, experience, and a job; some choose to make serving in the military their career.

While similar options will remain for the graduates of the future, marked differences are anticipated as well. Current projections suggest that enrollment in degree-granting postsecondary institutions will increase by 2026\(^2\), whereas participation in the work force among 16-24-year old Americans will decline during a similar time frame\(^3\). The lines between postsecondary education and work may also become blurred, as universities and corporations expand their partnerships\(^4\) to offer new educational opportunities and employers begin to develop their own training and credentialing programs.\(^5\)

Students will also likely continue to move in and out of, and between, postsecondary institutions\(^6\) and work experiences; the curve and length of individual pathways will reflect personal needs, interests, and goals. And though they will vary in their timelines, each pathway will eventually lead to some form of work and for most a career, typically as part of the paid labor force. But what will the work of the future look like?

The graduates of 2030 will need to be prepared for a postsecondary experience that may look very different from that for which their parents and teachers were prepared. Although jobs that sound quite familiar to us now (e.g., teacher, veterinary technician, lawyer, engineer) will likely still be widely available, it is impossible to predict the range of specific jobs that will be available and sought-after more than a decade from now. Trends such as globalization, automation, and “big data” point to major changes in the world of work, both in terms of the available jobs and the work environment in which those jobs will be carried out.

---


Technology linking employers and employees will allow geographically dispersed people to vie for the same job, leading to an environment in which American workers will compete with others from around the globe. Once hired, they will be part of an increasingly dispersed and diverse work force. Advancements in communications technology will continue to bridge this distance, allowing for collaboration across space and time. Work of the future will often be contract-based, and workers for the future will need to be prepared to work as part of cross-cultural and mixed-discipline teams. With this increase in contract work, work environments will be more fluid, with individuals able to determine their own schedule and/or work space.

Automation will reduce the number of human workers needed to perform routine tasks in some fields, while creating new job descriptions in other fields and adding jobs directly related to the machines used to automate work. Many customer service and middle management positions are expected to disappear as increasing numbers of transactions are completed via automated functions, thereby reducing the need for cashiers, clerks, and similar service providers, as well as those who would supervise them. Jobs relating to transportation and logistics, office and administrative support, manufacturing, and service are also expected to decline due to increased automation. At the same time, increasing numbers of workers will rely on artificial intelligence to assist them in their jobs. Humans and machines will collaborate to make decisions in the future world of work. Humans will also be responsible for designing and servicing these automated technologies.

Large quantities of data collected in real time will create job opportunities in data management and analytics, leading to expanded opportunities for those who can analyze and mine these data into information. Data collection mechanisms will be seamlessly integrated into all parts of life, and jobs will require a broad range of employees to use data. Due to this unprecedented access to real-time data, organizations will change quickly to meet the demands of the markets in which they function.

Job descriptions of the future will not be organized around clearly defined job titles, but rather around accumulated skills and experiences. Employers will leverage data to identify employees

---

who possess the experiences and skills specific to the job at hand.\textsuperscript{14} Potential employees will market themselves as uniquely qualified for the specifics of a project via online talent platforms and social media. Careers will be built out of freelance arrangements or other such “gigs”.\textsuperscript{15} Even those who opt for a more traditional career will likely hold around 12 jobs,\textsuperscript{16} either within one organization or across several organizations, over the course of their life. This will necessitate a continuous process of education and training throughout these future workers’ careers.\textsuperscript{17} Employers of the future will likely demand and incentivize on-going and just-in-time skill development to meet changing workforce needs.

Changes in the world of work have implications for postsecondary education as well. The graduates of 2030 will vary in terms of their high school experiences. Some will leave high school with college credits, Associate’s degrees and/or industry certifications, equipped with academic and job-specific knowledge and skills. Some will have paid work experience under their belts, in addition to or in lieu of service learning or other unpaid work experiences. Many will weigh decisions regarding what pathway is the best fit given their financial situation, their family’s needs, and their own goals, interests, and perceived abilities. Postsecondary institutions will need to use available data to make appropriate placement decisions, and then offer individualized instruction and a variety of student-focused services and support structures to ensure that students persist and ultimately graduate with competency in specific and generalized skills. Further, postsecondary institutions will need to adapt to meet the needs of an agile workforce that will be seeking opportunities for lifelong learning through additional credentials or courses designed to build new, or enhance existing, job skills.

**Requisite Skills for Future Work – Skills for What? (Q2)**

With this vision of the work of the future in mind, the skills needed to thrive in such an environment become apparent. Foundational academic skills, such as literacy and numeracy will continue to be valued and valuable in the postsecondary world of 2030. However, postsecondary success will also require a range of other cognitive, interpersonal and intrapersonal skills, as well as essential life skills. Cognitive skills include facility with technology, digital skills, computational thinking, and statistical literacy. Interpersonal skills include problem solving, communication, collaboration and cultural sensitivity. Intrapersonal skills include time management and flexibility. Essential life skills include financial and health literacy, as well as citizenship skills.

**Cognitive Skills**

Facility with technology will be in demand for jobs at all levels. Routine tasks formerly performed by humans may be taken over by robots or other technological advancements\textsuperscript{18}, leaving the human worker to service the technology or collaborate with the technology to complete more

\textsuperscript{17} PwC. (2018). *Workforce of the future.*
\textsuperscript{18} Ibid.
complex tasks. Employers will continue to increasingly seek individuals with computational thinking and digital skills to interact with data and new and emerging technologies.\textsuperscript{19}

Basic digital skills will be essential for developing other skills. Education and on-the-job training will increasingly be delivered via digital platforms and will incorporate virtual reality simulations more frequently.\textsuperscript{20} More advanced digital skills include those required to create and use digital tools. The creation of digital tools, such as artificial intelligence and machine learning, requires STEM, analytic, and computational thinking skills. Employers will need programmers and innovators to develop new technologies to tackle more difficult challenges and improve collaboration, efficiency, and cost effectiveness.

Data management and analysis skills will also be in demand across job categories as data become more accessible and easier to distribute and share.\textsuperscript{21} Computational thinking and statistical literacy skills will be sought after as employers seek those who can effectively use, visualize and manipulate, and draw conclusions from data. Statistical reasoning skills will also be in demand as new technologies require humans who are able to train machine learning algorithms, explain how they work, and keep them operating.\textsuperscript{22}

\textbf{Interpersonal Skills}
Regardless of the specific technical skill, workers of the future will not only need to possess such skills, but also will need to effectively apply those skills to real world problems. Problem-solving is an essential preparedness skill, one that will be required in more than one-third of jobs by the time the class of 2030 graduates high school.\textsuperscript{23} Employers and postsecondary educational institutions will seek graduates who are able to identify and select among many courses of action, and to do so in a dynamic work environment. Persistence in the face of uncertainty, and the ability to handle failure and identify next steps when things don’t go as planned, will also be essential for future postsecondary success.

Working as part of the dispersed and diverse workforce of the future will require communication and collaboration skills. Communication in the future world of work will require the ability to work with emerging communication technologies, along with the more traditional elements of communication such as listening and engaging in conversation. Active listening, in particular, is sought after by potential employers as it helps to create a positive work culture and supports collaboration, which in turn spurs innovation.\textsuperscript{24} Conversation skills are important because they...

\textsuperscript{23} Thompson, C. (2016). The top 10 skills that will be in demand by all employers by 2020. Retrieved from https://www.businessinsider.com/wef-report-skills-workers-need-2016-1
contribute to an organization’s shared understandings, which may be critical for agile decision-making. 25

Working collaboratively to solve problems will be an essential skill of the future, and will require building relationships, maintaining sensitivity to cultural differences, and seeing others’ perspectives. Forging positive relationships in the workplace will be key for an individual’s job satisfaction and will be an essential building block in the creation of a productive work environment. In the context of the workplace, cultural sensitivity includes working effectively alongside someone from a different cultural background who may approach workplace behaviors differently. Coworkers from different cultural backgrounds may engage in different behaviors and hold different work-related values, and culture-based misinterpretations can have implications for the success of collaborative efforts. 26 Perspective taking involves awareness of one’s emotions and empathy, and correctly interpreting what others are trying to communicate.

**Intrapersonal Skills**

The fluidity of work arrangements in the future will necessitate time management skills, as individuals work on multiple project teams or juggle multiple gigs. Time management requires a variety of skills: estimation of effort, scheduling, prioritizing, delegation, and monitoring a to-do list, among myriad others. An individual with strong time management skills can not only project the amount of time and effort a given task will require, but also inhabit the mindset to meet deadlines and, perhaps as importantly, recognize when a deadline cannot be met and adapt accordingly. While employers have historically valued employees with solid time management skills, in the expanding gig economy the individual entrepreneur’s personal success will depend upon it.

Flexibility and adaptability will also be essential skills, whether as an independent contractor providing services to multiple clients or as a career employee adjusting to changing roles and expectations in a dynamic environment. In fact, flexibility and adaptability, particularly in the context of interpersonal communication, are among the uniquely human skills that may prevent some occupations from becoming fully automated. 27 In addition to being key to specific careers, flexibility will also be integral to managing the projected evolving career path an individual will undertake over the course of a working lifetime. As companies demand upskilling or reskilling, the adaptable employee will be at an advantage.

**Life Skills**

Finally, graduating high school students will be best prepared for postsecondary success if they have acquired essential life skills. Skills such as financial literacy, and health and wellness literacy will help to ensure that decisions are made with long term implications in mind, setting the stage for continued success. Finally, high school graduates of 2030 will require citizenship skills, understanding the relations between individuals and society, the organization of the state

---


and how democracy functions, and the roles and responsibilities of the individual in the world around them.  

**Measures of Preparedness – Measures for what? (Q3)**

Postsecondary preparedness is a culmination of in-school and out-of-school experiences leading up to high school graduation. It is a multi-faceted concept that will require a multidimensional measurement approach. Currently, there is no uniform and comprehensive measure of postsecondary preparedness. NAEP measures academic preparedness for college without remediation with its reading and math assessments, but this does not address the wide range of skills discussed above. Existing measures such as NAEP assessments of academic skills will need to be combined with other existing data sources and will likely need to draw on new data sources or innovative measurement approaches, to gain a full picture of the state of postsecondary preparedness.

NAEP assesses a wide range of content areas at grade 12, including civics, economics, foreign language, geography, math, reading, science, technology and engineering literacy, U.S. history, and writing. The NAEP civics frameworks, for example, outline “intellectual skills” such as analyzing and evaluating, taking, and defending positions, along with participatory skills such as interacting, monitoring, and influencing. These intellectual skills may align with several interpersonal skills outlined above (e.g., communication, collaboration). Similarly, the technology and engineering literacy assessment frameworks outline three practices (understanding technological principles, developing solutions and achieving goals, and communicating and collaborating) that may reflect some of the described technical, intrapersonal, and interpersonal skills. NAEP also administers complex and interactive item types that may tap into skills beyond the academic skills measured. For example, a reading item may tap into cultural awareness skills, or a writing item may tap into effective communication skills. Finally, students who participate in NAEP are also encouraged to complete a student survey, which collects information about students’ educational experiences.

NAEP also collects data as part of the High School Transcript Study. This study is designed to reflect a nationally representative sample of U.S. schools and a sample of students that is representative of the graduates from each participating school. Data collected include the courses taken during high school, high school credits earned (including those earned in middle school), and final course grades. Information about course-taking patterns are not only informative about academic skills to which students have been exposed, but also provide indirect information about other important skills such as persistence and time-management.

Other data sources outside of NAEP, such as data regularly collected by or stored at state education agencies, may also be useful in informing our understanding of postsecondary preparedness. Participation in education and work experiences, from course attendance, to maintaining employment, to engaging in volunteerism or service learning experiences can

---

provide evidence of individuals’ development of important preparedness skills that can be aggregated to higher levels. Data embedded in these education and work experiences are another potential source of preparedness data. For example, experiences with technology, individual portfolios, or team projects are elements of applied educational and work experiences that yield information about relevant skills.

Data from assessments administered at the state level could also be mined for relevant information. For example, performance tasks designed to reflect real-world contexts may provide insight into blends of skills, such as problem-solving within an academic content area. Computer-based assessments could be a source of data on digital skills. Performance-based assessments may tap into persistence, problem solving skills, and analytical thinking skills.

New and emerging technological resources present additional opportunities to gather data from education and work experiences. Micro-credentials, also known as digital badges, are earned upon completion of a short course, administered online or in a more traditional classroom setting. In the future, blockchain technology may allow parties to record transactions and maintain a permanent digital record of them, which could be used to document achievements within a particular organization or program.\footnote{Iansiti, M. & Lakhani, K.R. (2017). \emph{The truth about blockchain}. Retrieved from https://hbr.org/2017/01/the-truth-about-blockchain}

**Current Context/National Need**

NAEP has been a leader in understanding and communicating what the nation’s students know and can do for decades. It has recently been argued that the timing is right for expansion of 12th grade NAEP testing, citing the need for trustworthy data on college and career readiness, and for a data source that would allow states to compare themselves to one another and to the country as a whole.\footnote{Finn, C.E. (2018). \emph{Time for twelfth grade state-level NAEP}. Retrieved from https://edexcellence.net/articles/time-for-twelfth-grade-state-level-naep} With its well-established infrastructure, NAEP is in a unique position to gather postsecondary preparedness data at the state level. NAEP data, along with other NCES data sources, will inform those interested in postsecondary preparedness policy, and will be useful to states as they work with industry and business partners to expand job opportunities, and to support local-level efforts to prepare students for life after high school. These data will also be of use to postsecondary education providers and employers as they seek to ameliorate gaps in preparedness, and to inform the public about how prepared high school graduates are for their next steps.

Over half of states have developed definitions of \emph{college and career readiness}.\footnote{American Institutes for Research. (2014). \emph{Overview: State definitions of college and career readiness}. Retrieved from https://ccrscenter.org/sites/default/files/CCRS%20Definitions%20Brief_REV_1.pdf} In a large majority of these states, a single definition is used to describe readiness for college and readiness for career, and some address preparedness for postsecondary life more broadly. Some of these definitions focus on academic skills necessary to enter credit-bearing college courses without remediation or to successfully complete postsecondary job training programs that will lead to a career that can support a family. Others describe additional skills beyond the academic, such as communication and collaboration. A very small number also include activities such as lifelong learning and civic engagement. Although there has been a substantial amount of work done at the state level, there are still states that have not developed a formal definition. Further, the existing state-level definitions are varied enough to make comparisons among the states confusing. NAEP is in a unique position to lead the charge in collecting data to inform a
common understanding and, by extension, identifying the most essential elements of postsecondary preparedness.
Appendix D. Expert Panel: Industry
Notes of the Expert Panel Meeting Representing Industry  
February 22, 2018  
National Assessment Governing Board  
Ad Hoc Committee on Measures of Postsecondary Preparedness

As part of meeting the charge of the Ad Hoc Committee on Measures of Postsecondary Preparedness, HumRRO organized and facilitated a meeting with industry experts. The purpose of this meeting was to get input from leaders and experts in industry about (a) the jobs that will exist in 2030, (b) the skills that these jobs will require, and (c) the measures/indicators that would be needed to provide a status of elementary and secondary students with respect to these skills.

We were fortunate to assemble an exceptional panel of experts and leaders. The panel members included Ms. Paula Collins, Texas Instruments, Mr. Marcelino Ford-Livene, Intel Corporation, Dr. Scott Heimlich, Amgen Foundation, Dr. Chauncy Lennon, JPMorgan Chase, and Mr. Reginald McGregor, Rolls-Royce Corporation.

The meeting was held on February 22, 2018 in Alexandria, Virginia. An overview of the National Assessment Governing Board and the charge of the Ad Hoc Committee on Measures of Postsecondary Preparedness, along with the agenda and logistical information for the meeting were sent to the panelists in advance.

Thanos Patelis (HumRRO) opened the meeting and after quickly informing the group of some logistics, Terry Mazany provided an overview and led the attendees through introductions. Then, Thanos Patelis facilitated the meeting around the three areas of inquiry involving (a) the jobs of 2030, (b) the skills that they will require, and (c) the measures/indicators that will be important to provide. Finally, Terry Mazany offered some concluding comments. The agenda and the list of all attendees is in Appendix A.

The purpose of this document is to provide information on the themes and comments made by the panelists. The information in this report is meant to provide insight into the rich conversation and comments provided by the expert panelists.

**The Future of the Workplace and Work**

- The titles of the jobs in 2030 cannot be predicted. However, the jobs of the future will require many skills and will be driven by globalization, artificial intelligence, and “big data”.
  - Globalization will change the workplace, from the types of jobs available (i.e., global competition for jobs) to working on cross-cultural teams.
  - Workplace integration will increase (e.g., working across disciplines instead of in silos by discipline).
  - The pace of automation and existence of the internet enable rapid access to information which will affect what employees do on the job and their job descriptions. The use of the internet and automation will only increase.
  - Employers should embrace new methods of communication, driven by the next generation. For example, hiring managers may not be familiar or may be uncomfortable with the latest communication modes of those applying for jobs. Rather than allowing that to impact negatively on job applicants, employers should
acknowledge the differences as innovation or trends to monitor. Job applicants may also need to be attuned to this dynamic.

- Technology will be at the forefront. For example, JP Morgan Chase is a “tech company that also loans money”; they do not consider themselves primarily a financial institution.
- Complicated tasks can be handled by automation (which will replace some jobs). Employees of the future will need to work with automated equipment and employees will be needed to design and service the automation.
- Complex tasks will take human thought (and these types of jobs will remain and additional ones will be added in the future).

- There is and likely there will continue to be a duality in the job descriptions of the future: academic skills and college degree required versus high school diploma and training and apprenticeship experience required. Panelists noted they come from the academic skills track and although they acknowledge the diploma-training track, they suggested consulting with experts in that area for a more detailed picture of what the future holds for those not following the 4-year college track.
  - Need to hire the person with the right skill set, not the person with the most qualifications (who may be overqualified and a poor fit for the work). This is sometimes a tendency when college-graduate hiring managers put more emphasis on college degree, the background they come from and perspective they bring to their job, than is warranted by the demands of the job being filled.
  - Most jobs that do not require a 4-year college degree, will require additional training, such as a 2-year college degree, technical training, or post-secondary education and/or training leading to certification.
  - Employer provides job skills (e.g., specific knowledge and procedures), while employee brings workplace competencies to the job (see competencies in the skills needed in the future). More job-related training will be provided by the employer, such as in-house mini-MBA programs provided by large corporations.
  - Continuous learning will be required to keep up with change. The employer will support or provide the training or education; the employee must participate to keep pace.

- Panelists indicated the need for initiatives to empower students, especially those who are “at-risk” and do not have role models, with an understanding of the labor market and expose them to employment options. Suggestions for empowering students so they are ready for post-secondary steps to meet their goals:
  - Help them define pathways to jobs.
  - Assist in setting goals; define an individual’s “north star”.

- Employer/employee relationships will change.
  - More contract work will emerge, which allows workers to dictate own schedule and/or workplace.

- Office space will be different.
  - For example, if employees come to the office, they will use a laptop and choose a work space area plugging into the network. The exact location may vary and will be more fluid than today.
Skills Needed in the Future

- Panelists described the need for employees to be able to apply skills, which defines competencies. Having a skill is not sufficient. Must know how to apply the skill to real world problems.
- The skills that were highlighted were as follows:
  - Ability to collaborate with people and machines, as the workplace incorporates more technology and automation as well as more collaboration.
  - Ability to interact with technology in jobs at all levels. Career Technical Education (CTE) can provide skills and certification for certain jobs.
  - Data skills are in demand - *data is the new oil*.
  - Less focus on job-specific content skills and more on workplace competencies:
    - Critical thinking, effective communication, collaboration, adaptability, problem solving, creativity, integrity, community/workplace citizenship, agility, learning disposition, persistence, attitude, interest.
  - Able to handle failure – *know what to do when the button fails*.
- Need power skills and experience, especially for at-risk students, to navigate the job market and succeed in entry-level positions – resume writing, oral communication, working on teams, basic reading/writing and mathematics ability.

Measures of Skills in the Future

- Consider measuring post-secondary readiness skills in grade 8.
- Maintain traditional knowledge measures (i.e., reading, mathematics).
  - Some went as far as to say that these measures of academic skills should not be removed and any other measures should be added.
- Design-build skills can be measured by persistence. Do you persist until object is built?
- Measure *application* of skills at grade 12. Can students demonstrate their skills (versus showing their knowledge of skills)?
- Add new measures tapping workplace requirements. Be creative in measuring skills (e.g., use certificates or credentials). Leverage CTE curriculum and measures.
  - In the interview process for candidates, hiring managers will give a problem to solve. Therefore, such metrics that demonstrate process and results of solving problems would be helpful.
- Need measures on collaboration, empowerment, and creativity.
- Tie relevancy of measures to industry and align with education. Do this regionally so that measures of preparedness are informative to:
  - students (do they have the skills needed for jobs in their community?),
  - industry (do local job applicants have the skills needed for jobs being offered in their community?),
  - educators (are they preparing students for post-secondary opportunities in their community?), and
  - policy makers (does the local workforce have the skills that industry in their community require?).
- While this may not be the Governing Board’s responsibility, students should be given the ability to develop digital portfolios, including coursework and experiential activities, in school to demonstrate their skills and achievements. This would be helpful to employers.
- The measures must keep evolving as the type of work and required skills change over time.
One interesting observation was that the panelists described job training interventions for at-risk youth with measures of program success embedded as artifacts of the experience. Did the participant build something? While the final product might not have been their initial design, the focus was on the creative process and the ability to troubleshoot problems as well as to persist in developing the final product.
Appendix A: Meeting Agenda and Attendees
Expert Panel Meeting
National Assessment Governing Board
Ad Hoc Committee on Measures of Postsecondary Preparedness

February 22, 2018 | Agenda

11:00 to 11:05 AM  Start Meeting
Thanos Patelis, Facilitator, HumRRO

11:05 to 11:15 AM  Welcome and Introductions
Terry Mazany, National Assessment Governing Board Member
Chair, Ad Hoc Committee on Measures of Postsecondary Preparedness

11:15 AM to 12:00 PM  Work of the Future
Thanos Patelis, Facilitator, HumRRO
Guiding Questions:
➢ What do you see as the type of jobs graduating high school seniors will have in 2030?
➢ Compared to jobs now, what kind of trends do you see emerging for jobs in 2030?
➢ Do you foresee any differences of jobs by industry or do you expect similar trends to occur for all jobs?
➢ What do you see as expectations of employers for these students?
➢ How do you envision the hiring process to be?
➢ What role will postsecondary institutions play in training and preparing students for these jobs?

12:00 to 12:15 PM  Break to get lunch

12:15 to 1:00 PM  Skills for the Work of the Future
Thanos Patelis, Facilitator, HumRRO
Guiding Questions:
➢ What types of skills will graduating high school seniors need to have in 2030 in order to get the jobs in 2030?
➢ What would you consider pre-requisite skills vs. skills that can be acquired on the job?
➢ What role will postsecondary institutions play in training these skills?
➢ What would a hiring manager in 2030 look for in prospective hires?

1:00 to 1:45 PM  Measures of these Skills Associated with Work of the Future
Thanos Patelis, Facilitator, HumRRO
Guiding Questions:
➢ What measures do you see being used to represent these skills?
➢ What metrics would provide helpful information in the aggregate about the skills of graduating high school seniors?

1:45 to 2:00 PM  Final thoughts and concluding remarks
Terry Mazany, National Assessment Governing Board Member
Chair, Ad Hoc Committee on Measures of Postsecondary Preparedness
Attendees

Expert Panelists:
- Paula Collins, Texas Instruments
- Marcelino Ford-Livene, Intel Corporation
- Scott Heimlich, Amgen Foundation
- Chauncy Lennon, JPMorgan Chase
- Reginald McGregor, Rolls-Royce Corporation

Governing Board Members:
- Terry Mazany, Chair, Ad Hoc Committee on Measures of Postsecondary Preparedness
- Honorable James E. Geringer, Former Governor of Wyoming, Cheyenne, Wyoming
- Carol Jago, Associate Director, California Reading & Literature Project at UCLA, Oak Park, Illinois
- Dale Nowlin, Teacher and Mathematics Department Chair, Bartholomew Consolidated School Corporation, Columbus, Indiana
- Honorable Beverly Perdue, Former Governor of North Carolina, New Bern, North Carolina
- Linda P. Rosen, Chief Executive Officer, Change the Equation, Washington, DC
- Chasidy White, Director of Strategic Initiatives, Office of the Superintendent, Montgomery, Alabama

Governing Board Staff Members:
- Bill Bushaw, Executive Director
- Lisa Stooksberry, Deputy Executive Director
- Lily Clark, Assistant Director for Policy & Research
- Laura LoGerfo, Assistant Director for Reporting & Analysis
- Munira Mwalimu, Executive Officer & Contracting Officer
- Sharyn Rosenberg, Assistant Director for Psychometrics
- Angela Scott, Management & Program Analyst

HumRRO Staff Members:
- Monica Gribben, Senior Staff Scientist
- Deirdre Knapp, Vice President, Assessment and Evaluation in Education and the Workplace
- Jackson Millard, Research Associate
- Thanos Patelis, Principal Scientist
Expert Panelists

Paula Collins
Vice President, Worldwide Government Relations
Texas Instruments

Paula J. Collins is vice president of Worldwide Government Relations for Texas Instruments where she leads the Company’s advocacy activities in the United States and abroad. She joined Texas Instruments in 1999 as Director of Government Relations and managed the Company’s legislative and public policy activities on a wide range of issues, including immigration, funding for basic research and education.

Ms. Collins came to Texas Instruments with extensive government, corporate and business association experience. After serving as a legislative assistant on Capitol Hill, she joined American Express Company, where for ten years she directed the Company’s legislative activities on a wide range of public policy issues including a number of trade initiatives. In 1993, she joined the Business Roundtable where she worked closely with corporate leaders to develop and implement public policy campaigns on international trade, budget and workforce initiatives. From 1995-1997, she directed international trade relations at Eastman Kodak Company and from 1997-1999 was a principal with The Fratelli Group, a strategic communications firm where she played an active role in the development and implementation of comprehensive public affairs strategies for several coalitions on trade and telecommunications issues.

Ms. Collins is a graduate of Yale University and attended the Program for Management Development at Harvard Business School. She is an active participant in her church and local civic organizations, and is a member of several professional organizations. She is a member of the Board of Directors and Executive Committee of the Information Technology Industry Council, and chairman of the Board of the Task Force on American Innovation.
Marcelino Ford-Livene
General Manager, Global Programs and Alliances
Intel Corporation

Marcelino Ford-Livene is the General Manager of Global Programs and Alliances for Intel’s Worldwide Corporate Affairs Group. In this capacity, he leads the organization charged with designing the framework and strategic plan for identifying and prioritizing win-win strategic alliances, relationships and partnerships with various global industry, government and special interest groups that advance the strategic direction of Intel’s Diversity and Inclusion Initiative. Prior to this role, Ford-Livene was the General Manger of New Channels and Advanced Advertising for Intel Media, where he led the organization charged with programming, licensing and distributing new format television channels and advertising-supported video-on-demand programming. He was also responsible for advertising sales, advertising operations, audience research and data analytics for Intel Media’s OTT services. He also co-authored patents on TV viewership analytics and advanced advertising behavioral targeting. Prior to Intel, he was a senior member of TV Guide’s corporate development and planning team. He has also held senior positions with the U.S. Federal Communications Commission in Washington, DC. He served as Special Counsel for New Media Policy for Chairman William E. Kennard and as Senior Counsel and Director of Media Strategic Analysis for the FCC’s Office of Strategic Planning under Chairman Michael Powell. Ford-Livene was the Division Chairman of the Interactive Media Division for the American Bar Association’s Forum on the Entertainment and Sports Industries from 2006 to 2013. He also served for eight years on the board of the TV Academy, the organization that awards the prestigious Primetime Emmy for creative excellence in the television industry. He was also the TV Academy’s Board Secretary and a member of its Executive Committee from 2010 to 2013. He is currently the Co-Chairman of the TV Academy’s Diversity Committee and a founding board member of the Digital Diversity Network. Corporate boards that Ford-Livene has served on include Delivery Agent in San Francisco, CA and TRA Global, which was acquired by TiVo. Ford-Livene earned a B.A. in economics from UC San Diego, a J.D./M.B.A. from the University of Illinois and has completed an Executive Leadership Program at Harvard Business School.
Scott Heimlich
Vice President, Amgen Foundation

Scott M. Heimlich is vice president of the Amgen Foundation. He is responsible for the strategic management and direction of the Foundation’s science education portfolio, including the development and oversight of key initiatives at the K-12 and higher education levels. He was the principal architect and continues to lead the Amgen Scholars Program, the Foundation’s largest initiative providing undergraduates with access to research opportunities at premier educational and research institutions across the world. Under his leadership, the Amgen Biotech Experience transformed from a local program into a multi-site, international initiative bringing biotechnology lab experiences to over 80,000 secondary students a year. With these and many other initiatives, the Foundation’s commitment to science education recently surpassed the $125 million milestone.

Prior to joining Amgen in 2005, he served in positions at the University of California, Los Angeles, Los Angeles Pierce College, University of Southern California, and a junior high school in Japan. He holds a bachelor’s degree, master’s degree, and doctorate in education from the University of California, Los Angeles.

Chauncy Lennon
Managing Director and Head of Workforce Initiatives
Global Philanthropy
JPMorgan Chase & Co.

Chauncy Lennon leads JPMorgan Chase & Co.’s initiatives to promote economic opportunity through investments in workforce practice, innovation, and policy. These include New Skills at Work, a $250 million global initiative to support demand-driven workforce systems that promote prosperity for workers and industries; New Skills for Youth, a $75 million initiative to increase the number of young people who complete career pathways that begin in high school and end with postsecondary degrees or credentials aligned with good-paying, high-demand jobs; The Fellowship Initiative, a program providing young men of color with learning experiences that help them achieve their education and career potential; and a $17 million investment in Summer Youth Employment Programs in US cities to help underserved youth obtain the skills necessary to build lasting careers.

He serves on the New York City Workforce Development Board, the College Promise Campaign Advisory Board, and the Neighborhood Trust Financial Partners Board.
He joined JPMorgan Chase from the Ford Foundation, where his grant-making focused on promoting economic advancement for low-income workers by improving access to workforce development and work support programs. Prior to the Ford Foundation, he was senior vice president for Asset Building at Seedco, a national workforce development intermediary. He also has extensive experience researching the mobility patterns of the working poor. He earned his Ph.D. in anthropology from Columbia University, master's degree from the University of Chicago and bachelor's degree from Williams College. He has taught urban studies at Columbia's School of International and Public Affairs and Barnard College.

**Reginald McGregor**  
Manager, Research & Technology Strategy Group  
Rolls-Royce Corporation

Reginald McGregor, Manager of Engineering Employee Development and STEM Outreach at Rolls-Royce Corporation. He is a Mechanical Engineer with over 15 years' experience in various engineering roles. He spent over 8 years in early career development managing the engineering co-op; high school internship and graduate development programs. Reginald holds BS in Mechanical Engineering, MBA and currently completing a MS in Technology Leadership and Innovation. He is very active in workforce development and STEM education and serving the community. Reginald enjoys reading, outdoor activities and spending time with family.

Reginald serves on several boards and committees including the Governor-appointed Region 5 Works Council, President of the Lawrence Township School Board, Indiana STEM Advisory Council, STEMx National Advisory Board, Purdue Engineering Education Industrial Advisory Council, Marion County Superintendents STEM Coalition, Indiana Chamber of Commerce K-12 and Workforce Committees, Million Women Mentor Steering Committee, Indiana Afterschool Network Board, and EmployIndy Youth Committee.
Appendix E. Expert Panel: Higher Education
As one step in addressing the charge of the Ad Hoc Committee on Measures of Postsecondary Preparedness, HumRRO organized and facilitated a meeting with a select group of higher education innovators. The purpose of this meeting was to elicit input from leaders and experts in higher education about (a) the jobs that will exist in 2030, (b) the skills that these jobs will require, and (c) the measures/indicators that would be needed to determine the status of elementary and secondary students with respect to these skills.

We were fortunate to assemble an exceptional panel of experts and leaders. The panel members included Dr. Sarah DeMark, Vice President of Academic Programs, Western Governors University; Dr. Pradeep Kotamraju, Bureau Chief, Career and Technical Education, Division of Community Colleges and Workforce Preparation, Iowa Department of Education; Mr. Michael Morsches, Dean of Learning Enrichment and College Readiness, Moraine Valley Community College; Dr. Yvette Mozie-Ross, Vice Provost for Enrollment Management and Planning, University of Maryland, Baltimore County; and Dr. Holly Zanville, Senior Advisor for Credentialing and Workforce Development, Lumina Foundation. Also, in attendance were some Governing Board members, Governing Board staff members, and HumRRO staff, listed in Appendix A.

The meeting was held on April 19, 2018 in Chicago, Illinois. An overview of the National Assessment Governing Board and the charge of the Ad Hoc Committee on Measures of Postsecondary Preparedness, along with the agenda and logistical information for the meeting were sent to the panelists in advance of the meeting.

Thanos Patelis (HumRRO) opened the meeting and after quickly informing the group of some logistics, Terry Mazany, Ad Hoc Committee Chair, set the stage for the role of NAEP in the future, given the impact of technology on work as well as the economic and global context in which students enter the post-secondary world. He led the attendees through introductions. Thanos Patelis facilitated the meeting around the three areas of inquiry involving (a) the jobs of 2030, (b) the skills these jobs will require, and (c) the measures/indicators needed to measure these skills. Finally, Terry Mazany offered some concluding comments. The agenda and the list of all attendees is in Appendix A.

The purpose of this document is to summarize the themes and comments made by the panelists. The information in this report is meant to provide insight into the rich conversation and comments provided by the expert panelists.
The Future of the Workplace and Work

With experts representing higher education, the discussion of the future of the workplace and work focused on pathways to work, primarily through postsecondary education and training.

- Postsecondary institutions need to create pathways to develop agile employees who are open to lifelong learning.
- Lifetime or continuous learning will become the norm. Employees will need to continue to learn from different providers, from colleges/universities to specific training courses to experiential opportunities, throughout their lives. Information technology (IT) workers already face this with a variety of certifications for specific technology tools and applications. Highly-regulated occupations will likely be the last ones to make changes.
- Postsecondary institutions need to partner with employers to identify education and training needs so that graduates possess the knowledge and skills needed for jobs.
  - Look to IT which is leading the way in defining job requirements and credentials for employees.
  - One of the panelists described a keynote presentation by the CEO from Chegg, Dan Rosensweig, describing the current disconnect between expectations and responsibilities of employers, higher education, and students. He illustrated this by placing each of the stakeholders at the vertices of a triangle with arrows facing outward indicating a lack of working together rather than arrows pointing inward, toward each other, signaling collaborative planning and working together toward similar goals.
  - Educators can be resistant to business models.
- There are still barriers to postsecondary education. Although community colleges have an open policy (in some states students do not need a high school diploma to enroll in community college), students may find it difficult to pursue their desired major or to matriculate. Prerequisites and competitive admission in selected programs (e.g., healthcare) are barriers to entry.
  - Similarly, some 4-year colleges guarantee admission to those with associate’s degrees, but cannot guarantee admission into specific programs due to enrollment capacity and accreditation requirements such as completion of specific coursework.
  - Some community college graduates are not prepared for 4-year colleges and universities because their 2-year institutions have limited qualification requirements for instructors and low standards for their graduates. Both of these factors could be a barrier to continued education.
- More individualization in postsecondary education requires “policy by anomaly.”
  - In developmental education, need to identify what students need and how to get it to them. Placing students on paths matching their goals raises retention rates.
- Strong partnerships are needed between 2- and 4-year institutions of higher education to facilitate students’ transfer between schools.
  - High school graduation projections show Hispanics are the fastest growing group\(^{36}\) and many of this group begin their postsecondary studies in community college.
  - Many students are graduating from high school with associate’s degrees obtained through early middle college programs and dual enrollment.
- Colleges and universities must provide different, perhaps individualized, services to students who enter at different points on the pathway to a 4-year degree. Historically, 18-year-old high school graduates enter as freshmen with new-student services and support structure

for the first year or two. Institutions are now called on to help a select group of high school graduates entering college with associate’s degrees, yet perhaps still needing wraparound services due to their youth (compared to the services offered to 20-year-old or older students transferring to a 4-year program with an associate’s degree). Other students may start and stop their education multiple times and attend several institutions before graduating.

- To prepare students for future jobs, we need vertical and horizontal articulation. For horizontal articulation, students need technical, academic, and employability skills (e.g., grit, self-understanding). For vertical articulation, the key is determining at what age/grade to start. High school staff say it needs to start in middle school; middle school staff say it needs to start in elementary school.
- Need a mechanism to validate training and experience as part of the pathway to a degree. More and more high school graduates are already working through the gig economy. Other students have jobs and families while attending college.
  - Look to the military; they validate training as credits.
  - Western Governors University (WGU) provides micro-credentials or badges as students achieve milestones to show them the skills and knowledge attained as they work toward their bachelor’s degree.
  - Give students the ability to curate their work and educational experiences.
- There is tension between an integrated approach providing a broad range of skills (academic, technical, and employment-oriented) and the business need for a narrow, specific set of skills to meet a skill shortage. One is too esoteric, the other too pragmatic.
- Post-secondary institutions will not be the destination, but a vehicle for certifying student competencies.
- Expect the acquisition and use for knowledge and skills to flip. Currently, knowledge is the base foundation provided by formal education and we obtain skills as needed. In the future, skills will be the base and we will obtain knowledge as needed.

**Skills Needed in the Future**

- Don’t teach students to do what a robot can do better.
  - Robots are better than humans at pattern recognition, repetitive tasks, etc. but they are not able to understand nuance of language, social relationships, or creativity.
  - It will be important for humans to connect domains.
  - McKinsey has developed a list of human skills such as empathy, planning, creativity, common sense, sense making, novel thinking, nuance of language, social relationships, etc.\(^{37}\)
- In addition to content or professional knowledge, students need:
  - practical transition skills
  - key learning skills and cognitive strategies
  - strong foundation of self-understanding and engagement strategies
  - critical thinking
  - affective mindset and skills
  - meta learning
  - financial literacy
  - information technology literacy
  - health and wellness literacy.
- Schools can provide learning and workplace skills.

---

• College experience courses for high school students.
• WGU offers eight synchronous online sessions with a small, facilitated cohort on skills such as self-efficacy, communication, and learning styles. In a pilot test with at-risk students, there were significant positive outcomes: performance in courses as well as retention increased. Some of the skills, including leadership and communication, were identified by the medical profession as ones missing in graduates. These skills not only make graduates better job candidates but also more resilient students.

• Consider where or why skills are needed to build awareness of how skills fit into work.
• Four-year institutions look for grit or persistence as a necessary skill for student success. Students with a solid academic foundation and grit should be able to succeed, whereas students with a strong foundation of academic knowledge and no grit may not be able to handle the rigor of college.
• Class attendance is the best predictor of success, as evidenced both by anecdote and research. Some colleges require attendance and initiate interventions if students do not attend class.
  o There is a question of how to measure attendance for online courses. One approach is to look at student engagement using interaction data from Learning Management Systems (LMS).
• Students need to learn how to get “unstuck” when in a challenging situation.
• Employers are looking for people who can work across left and right brains and are able to work with technology.

Measures of Skills in the Future

• Employers offer performance-based pay for high-value, high-priority credentials supporting ability to use skills.
  o Students may demonstrate their skills through portfolios.
  o Use blockchain to document achievements and portfolio.
• Need new types of student assessment.
  o Current assessments focus too much on knowledge and not enough on skills, character, and meta learning.
  o Students take most current assessments working alone rather than in teams. Need authentic assessments of team work with hands-on performance components.
• Leading-edge assessments use simulation and are more applied, with problem solving scenarios that assess whether you can use knowledge.
• Create dashboards for parents and students to see skill attainment, including credentials.
• Use micro credentials and then stack those credentials to meet employer-relevant needs.
• There is a tension between broad versus specific measurement of skills.
• Include all stakeholders in identifying what and how to measure skills.
• Measuring college or postsecondary readiness is different than college or postsecondary success.
• Some postsecondary institutions use transcripts, others don’t.
  o Transcripts could provide an opportunity to leverage high school data for postsecondary instructors to know what students have done prior to college and to personalize postsecondary instruction.
  o Expect seat time to be a less helpful measure from an industry perspective. They will be interested in a “transcript” with learning opportunities, perhaps using blockchain technology.

38 For information about blockchain: https://hbr.org/2017/01/the-truth-about-blockchain
For transcripts to be useful to instructors, need a way to standardize them.
- Need to include attendance on transcript.
- Metrics of academic rigor exist with validity evidence provided to support their value in predicting college outcomes.
- Concern with the shelf life of measures such as SAT or ACT, course grades, etc. Are high school results as valid for older, returning students?
- Metrics should include student employment.
- Measures of service learning are needed.

**Reflections**

Terry Mazany offered four reflections on the discussion:

1. We need to project all of the allied trends in society to 2030. Work is shifting to a gig economy. This will be the reality for 16- to 18-year-olds in 2030. We need to factor the expected changes in the economy of 2030 into the skills required to work in the future. Data is the new oil. Micro-credentialing and digital badges will more and more populate transcripts and portfolios.

2. There will be several paradigm shifts: (a) knowledge/skill flip, (b) everything has a developmental progression except technology, (c) the nontraditional student of today will be the traditional student of tomorrow, (d) students will be agents for themselves, and (e) a world where trust is collapsing in every venture except nonprofit ventures – blockchain as a key to build this trust.

3. We are in-between systems. We need to maintain an ecological perspective of each part of the system and look at the reciprocal changing role of employers.

4. The role of NAEP: We need to align NAEP with the requirements of Every Student Succeeds Act (ESSA), such as conditions of learning. This might be done by back-mapping the requirements of ESSA with what NAEP provides.
Appendix A: Meeting Agenda and Attendees
Expert Panel Meeting
National Assessment Governing Board
Ad Hoc Committee on Measures of Postsecondary Preparedness

April 19, 2018

11:00 to 11:05 AM
Start Meeting
Thanos Patelis, Facilitator, HumRRO

11:05 to 11:15 AM
Welcome and Introductions
Terry Mazany, National Assessment Governing Board Member
Chair, Ad Hoc Committee on Measures of Postsecondary Preparedness

11:15 AM to 12:00 PM
Work of the Future
Thanos Patelis

Guiding Questions:
- What do you see as the postsecondary pathways that high school seniors graduating in 2030 will be choosing among? (11:15-11:40)
- Compared to now, what kind of trends do you see shaping postsecondary education in 2030? (11:40-12:00)

12:00 to 12:15 PM
Break to get lunch

12:15 to 1:00 PM
Skills for the Work of the Future
Thanos Patelis

Guiding Questions:
- How have postsecondary entrance expectations changed in recent years? (12:15-12:40)
- What types of competencies and content knowledge will graduating high school seniors need to be prepared for postsecondary pathways in 2030? (12:40-1:00)

1:00 to 1:45 PM
Measures of these Skills
Thanos Patelis

Guiding Questions:
- What measures do you see being used for these competencies? What will require new or updated measurement tools? (1:00-1:20)
- What metrics would provide helpful information in the aggregate about the competencies of graduating high school seniors? (1:20-1:45)

1:45 to 2:00 PM
Final thoughts and concluding remarks
Terry Mazany
Attendees

Expert Panelists:
- Sarah DeMark, Vice President of Academic Programs, Western Governors University
- Pradeep Kotamraju, Bureau Chief, Career and Technical Education, Iowa Department of Education
- Michael Morsch, Dean of Learning Enrichment and College Readiness, Moraine Valley Community College
- Yvette Mozie-Ross, Vice Provost for Enrollment Management and Planning, University of Maryland, Baltimore County
- Holly Zanville, Senior Advisor for Credentialing and Workforce Development, Lumina Foundation

Governing Board Members:
- Terry Mazany, Chair, Ad Hoc Committee on Measures of Postsecondary Preparedness
- Dale Nowlin, Teacher and Mathematics Department Chair, Bartholomew Consolidated School Corporation, Columbus, Indiana
- Alice Peisch, Legislator, Massachusetts House of Representatives, Wellesley, Massachusetts
- Chasidy White, Director of Strategic Initiatives, Office of the Superintendent, Montgomery, Alabama

Governing Board Staff Members:
- Bill Bushaw, Executive Director
- Lisa Stooksberry, Deputy Executive Director
- Lily Clark, Assistant Director for Policy & Research

HumRRO Staff Members:
- Monica Gribben, Senior Staff Scientist
- Sunny Becker, Principal Staff Scientist
- Thanos Patelis, Principal Scientist
Expert Panelists

Sarah DeMark, Ph.D.
Vice President of Academic Programs
Western Governors University

Sarah DeMark joined nonprofit Western Governors University (WGU) in September 2014, and serves as the Vice President of Academic Programs, responsible for leading WGU’s portfolio strategy as well as the design and development of the university’s competency-based degrees, curriculum and assessments. This portfolio includes more than 50 programs, 600 courses, and nearly 1000 assessments.

Prior to joining WGU, DeMark spent more than 15 years at leading IT companies, serving in various leadership roles where she oversaw the strategy and execution of the design, development, and deployment of certification and curriculum-based assessment portfolios. Previously, she was an independent consultant working with state and local school districts, as well as working with The College Board on SAT and AP program evaluation.

DeMark is published in numerous journals and books and is a sought-after speaker. DeMark currently sits on ANSI’s Personnel Certification Accreditation Committee, which serves to validate whether certification programs adhere to standards.

DeMark earned a Ph.D. in Educational Psychology (Measurement, Statistics, & Methodological Studies) from Arizona State University. DeMark earned B.S. degrees in both Elementary Education and Psychology from Vanderbilt University.
Dr. Pradeep Kotamraju is currently the Bureau Chief, Career and Technical Education, Division of Community Colleges, Iowa Department of Education. As Iowa’s State Director for Career and Technical Education (CTE), he has leadership responsibility in managing those secondary and community college CTE programs that are funded through the Carl D. Perkins federal program. Previous to his current position as the Iowa CTE State Director, Dr. Pradeep Kotamraju has served the Deputy Director, National Research Center for Career and Technical Education (NRCCTE), University of Louisville, Louisville, Kentucky. Prior to that, he served as the System Director, Perkins, at the Minnesota State Colleges and Universities, Office of the Chancellor. Dr. Kotamraju has worked in several senior administrative positions in higher education and workforce development agencies in Minnesota.

Dr. Kotamraju has written several publications and monographs, and made numerous presentations, in the area of student success in career and technical education, workforce development in the United States, and, in the area of economic progress in the developing world. His research has included the examination of a variety of labor market information and workforce development issues that connect occupations, skills and careers, as individuals transitioned back and forth between employment and education. Dr. Kotamraju has been invited to participate on several statewide, regional and national committees that have focused on CTE programs, budget and finance, and accountability. Some of these committees have had even broader focus that places CTE right front and center when it comes to connecting education, workforce development, and economic development.

Before working in the public sector, Dr. Kotamraju taught college- and university-level Economics and Statistics at several higher education institutions in Minnesota and Kentucky. Dr. Kotamraju holds a Ph.D. in Economics from the University of Illinois. He received his Masters Degree in Economics from George Washington University, and his Bachelors in Economics from the University of Delhi, India.
Michael Morschès
Dean of Learning Enrichment and College Readiness
Moraine Valley Community College

Michael Morschès has worked in higher education for more than thirty years. His primary focus has been on developmental education and the transition from high school to college.

Michael currently serves as the Dean of Learning Enrichment and College Readiness at Moraine Valley Community College. He oversees the ABE/GED, ESL, developmental education, literacy volunteers, and tutoring programs. Michael has published numerous articles and handbooks on retention, student engagement, and teacher training in post-secondary institutions.
Yvette Mozie-Ross, Ph.D.
Vice Provost for Enrollment Management and Planning
University of Maryland, Baltimore County

Yvette Mozie-Ross, PhD, is Vice Provost for Enrollment Management and Planning at the University of Maryland, Baltimore County (UMBC). As Vice Provost, Dr. Mozie-Ross provides oversight and strategic planning for the areas of undergraduate admissions and orientation, financial aid and scholarships, academic and pre-professional advising, records and registration, and the student administration project (student information system). With a higher education career spanning over 25 years, she has served in numerous professional capacities including residence community director, coordinator of multicultural recruitment, assistant director for transfer recruitment and admissions, director of undergraduate admissions, and director of academic services (advising and registration). Dr. Mozie-Ross has served on various national and statewide committees and workgroups including the College Boards’ Commission for Transfer Policy and Practice, and the Maryland Higher Education Commission’s State Plan Writing Group on Access, Affordability and Completion. She has served on the university’s Strategic Planning Steering Committee and is currently serving as a member of the governing board for the Baltimore Collegetown Network, a consortium of 13 colleges in Baltimore, Maryland. Dr. Mozie-Ross frequently lends her expertise, both nationally and internationally, in the area of data analytics and leveraging analytics for institutional transformation. Dr. Mozie-Ross earned her bachelor’s degree from UMBC in 1988, her master’s degree from University of Maryland University College in 1994, and her doctorate in Education Policy and Leadership at the University of Maryland, College Park in 2011. Her dissertation research examined the academic and background characteristics of high school graduates who identified teachers as influential in their choice of college. Dr. Mozie-Ross enjoys spending time with her husband of 22 years and their 20-year old son. Her pass-time interests include family genealogical research and running.
Holly Zanville, Ph.D.
Senior Advisor for Credentialing and Workforce Development
at Lumina Foundation

Holly Zanville is Senior Advisor for Credentialing and Workforce Development at Lumina Foundation. She leads a new portfolio on Worker and Employer Engagement that focuses on building the capacity of educators and employers to scale and spread the best ideas in training, credentialing, and other workforce development strategies linked to postsecondary learning opportunities; and examining issues around the future of work and learning. Her work includes cultivation of networks and partnerships essential to the emerging new postsecondary learning system including Credential Engine, quality assurance efforts to ensure that credentials stand for high-quality learning, and networks for research and industry sector engagement. She previously led Lumina’s development of the national Connecting Credentials initiative, credential completion for returning adults with prior college/no credential, and statewide approaches to reverse-transfer degrees through the Credit When It’s Due initiative. Zanville received her Ph.D. in Educational Administration from the University of Minnesota; MA in English from the University of Wisconsin-Madison, and BA in English and Biology from Lindenwood University.
Appendix F. Expert Panel: Futurists
As one step in addressing the charge of the Ad Hoc Committee on Measures of Postsecondary Preparedness, HumRRO organized and facilitated a meeting with a select group of futurists. The purpose of this meeting was to elicit input from thought leaders regarding the future of postsecondary education and work.

We were fortunate to assemble an exceptional panel of visionaries with a variety of perspectives. The panel members included Randy Bennett, Educational Testing Service; Karen Cator, Digital Promise; David Conley, EdImagine; Alana Dunagan, Clayton Christensen Institute; Devin Fidler, Rethinkery Labs, and Nancy Lue, Advanced Education Research and Development Fund on behalf of the Chan Zuckerberg Initiative and the Bill & Melinda Gates Foundation. Also, in attendance were several Governing Board members, Governing Board staff members, and HumRRO staff.

The meeting was held on June 21, 2018 in San Francisco, California. An overview of the National Assessment Governing Board and the charge of the Ad Hoc Committee on Measures of Postsecondary Preparedness, a “facebook” of attendees with brief biographic summaries, along with the agenda and logistical information for the meeting were sent to the panelists in advance of the meeting. Appendix A contains the agenda, list of attendees, and panelist biographies.

Terry Mazany, Ad Hoc Committee Chair, welcomed the futurists and set the stage for the role of NAEP in the future, given the impact of technology on work as well as the economic and global context in which students enter the postsecondary world. He led the attendees through introductions. Thanos Patelis (HumRRO) reviewed the agenda and stated the goals for the meeting.

To establish the perspectives of these varied experts, each panelist provided a 10-minute presentation of their initial thoughts regarding five discussion questions: (a) what are the trends you see that will define the future of learning and schools? (b) what are the trends you see that will define the future of work and the skills that will be most valued by employers of the future? (c) what are the most promising technologies that will redefine education? (d) what things are most likely to disrupt how we think about teaching and learning? and (e) what are the trends that most concern you, and why? Copies of the presentation slides are in Appendix B.

Following the presentations, Thanos Patelis facilitated deeper discussion about common themes and the five questions. Finally, Terry Mazany offered some concluding comments.

The purpose of this document is to summarize the themes and comments made by the panelists. The information in this report is meant to provide insight into the rich conversation and comments provided by the expert panelists.

---

39 Although some panelists would not describe themselves as “futurists,” per se, their careers all include the identification and evaluation of trends, as well as forecasting future conditions or developments.
**Presentations**

**Randy Bennett** described seven trends in the future of learning.

- Learning is increasingly technology-based with complex tasks (e.g., simulation and games).
- Materials and methods used in learning are only now catching up with cognitive science.
- Learning is more person-based, adaptive, and customized on different dimensions, to (a) allow accessibility to make learning more available to students with diverse learning types, (b) personalize in terms of competency level, (c) engage students effectively, and (d) give students greater agency over their learning goals.
- New constructs and competencies, such as socioemotional learning, citizenship and citizen engagement, and cross-cultural competency, are becoming more prevalent.
- Prior knowledge is critical when learning new information or developing new skills.
- There is a focus on cross-disciplinary skills such as communication and problem solving. However, contextual differences within disciplines are important considerations (e.g., problem solving in art differs from problem solving in science).
- Assessment embedded in instruction with automated analysis and feedback, allows for adjustment of instruction.

In addition to trends in the future of learning, Dr. Bennett described two trends of most concern.

- Personalization – There is concern that personalization could be used to exacerbate as much as ameliorate differences in opportunities and learning. For example, students from underrepresented groups could be routed toward basic skills classes.
- Embedding assessment in instruction – There is potential for embedded assessment in instruction for student learning, however conflating assessment for learning with assessment for accountability could be problematic, especially if used to make policy judgements.

**Karen Cator** provided the following perspectives regarding the five questions:

- Trends in the future of learning include: (a) personalization to accommodate variability in students through learning science, (b) more flexible learning to obtain and demonstrate competency, and (c) performance-based assessments leading to credentials for the changing global workforce.
- Trends in the future of work and skills include artificial intelligence (AI) which has the potential to disrupt many jobs. Employees will need deeper learning skills such as collaboration and social emotional skills. We should focus on what is uniquely human.
- Technology can be used to augment human performance. For example, data from embedded assessment and improved diagnostics can provide more precise and accurate analyses of student knowledge and performance, helping teachers perform more effectively in the classroom.
- Learning science could be disruptive. People will have jagged profiles—different levels of competence across skills—based on individual differences and the contexts in which they apply the skills.
- Most concerning is disenfranchisement of teachers. As an example, one-third of current teaching jobs in St. Louis are vacant. Other areas of concern include limited resources in

---

40 Ms. Cator recommended Jack Ma’s presentation at the World Economic Forum on The Way We Teach; https://www.youtube.com/watch?v=pQCF3PtAaSg.
schools, increasing cost of higher education, limitations of current assessments, equity of access to quality learning activities, and the digital learning gap.

David Conley shared the following insights regarding the five questions:

- The future of learning includes the following trends: (a) taking the teacher out of the bottleneck role, thereby allowing students to work at their own pace and receive just-in-time learning; (b) providing more social learning; (c) using technology to identify learning patterns to personalize learning; and (d) focusing on adapting skills to accommodate changes in work rather than learning fixed skill sets.
- Trends in the future of work and skills include changes such as (a) gig work versus long-term careers, (b) continued adaptability, (c) hiring at low- and high-skill end with less at the middle-skill level, (d) global work teams while living locally, (e) increasing service work, and (f) standardization versus bespoke work (see jagged profiles as mentioned by Ms. Cator).
- Promising technologies in education are adaptability, including a wider variety of students, specialized job/task-specific reading, and web-based learning.
- The following may contribute to disruptions in teaching and learning: (a) students having more agency over their learning, (b) basic skills taught in context using simulations or serious games such as used in the military and medical training, (c) self-directed learning will require resources for teachers to help students who have trouble directing their own work, and (d) emphasis on career preparation with certifications and badges over liberal arts education.
- The three most concerning trends are (a) equity in education, (b) equity in defining preparedness, and (c) increasing the pace of disruptive economic change.

Alana Dunagan offered the following insights:

- The future of learning and work includes the following trends:
  - increased online learning in higher education and K-12
  - learning not requiring a terminal degree (e.g., a certification)
  - workforce alignment of education
- Higher education institutions are seeing falling enrollment, while training in specific skills matters more. She described the parallels between disruptive innovations in education and in business. She explained that corporate bankruptcy following implementation of disruptive technology occurs when companies do not adapt by using technology to expand the reach of their services (i.e., they continue serving the same set of customers rather than expanding their customer base); Blockbuster is an example of this situation.
- Jobs requiring higher education are growing twice as fast as jobs that do not, because of disruption by the education technology market. Innovators in the education technology space are developing partnerships with employers and creating new ways of offering higher education providing the needed training.
- The biggest concern in education and work is the prestige-based model of signaling competence (i.e., a degree from an elite university is highly valued over a degree from a lower tier school without regard to a student’s actual knowledge and skill). This model ignores the skills a student has and does not include employers in identifying the skills that students should learn. A better model would engage businesses in identifying skill needs, offer education aligned to workforce needs, and provide students with evidence of skill attainment and a means for submitting that information to employers.
Devin Fidler provided a historical perspective to inform the following trends:

- The history of change in organization strategies evolved from guilds to industrialization to manufacturing/assembly to digital. The advent of the World Wide Web facilitated communication and has expanded to commerce and coordination.
- Examples of using technology to speed up work include peer to peer applications such as TaskRabbit, Gigwalk, and Upwork. These platforms have millions of people enrolled to offer their services with qualifications based on past performance. Employers can use these applications to identify well-qualified candidates with the appropriate skills mix and a history of positive reviews; employees can use these applications to find jobs and to see what skills are in demand.
- The most promising technologies are using organization technologies in education technology with artificial intelligence.
- Disruption will come from small innovative organizations who are more nimble than large businesses.
- The biggest concern is the stereotype that organization is dehumanizing; however, organization can expand human capability.

Nancy Lue identified the following education megatrends:

- Return on education (i.e., value of education) is similar to an internal rate of return (IRR), a term very popular in Silicon Valley to evaluate whether to invest in something. In 2015, 50% of college graduates were working in fields they did not study in school, and more than a third indicated they would study different subjects if they had the opportunity to do it all over again. In a 2013 survey, nearly 40% of college graduates indicated university did not prepare them for employment. Meanwhile, companies and non-profit organizations (e.g., Coursera, EdX, Khan Academy) offer courses for free or with a credential for $100. General Assembly boasts a 98% success rate in securing jobs or promotions within six months for their graduates.
- Continuous learning (e.g., Kaizen education) offers an opportunity for adults to keep up with education. Millennials are projected to have 15 careers in their lifetimes so this ongoing education is important. Coursera’s MOOC “Learning How to Learn” has been highly successful, with over half a million “alumni.”
- Technology provides opportunities for ongoing learning. Video games are one venue for learning. One way to get people interested in education is to make the best (aka, “rock star”) teachers available through technology.
- Knowledge increasingly can be seen as currency (e.g., micro-credentials, badges). The degree-driven, prestige-driven education system isn’t meeting the needs of modern society. “What you know” is becoming more important than “where did you go.” Individuals can curate a portfolio of skills evidenced by micro-credentials, etc.
- Big data and smart data provide an opportunity to use data to personalize learning (Dreambox, Knewton, etc.).
- Mobile technology learning applications are widely available. Today, 90% of high school and college students own a smart phone, and time spent on smart phones is increasing. Smart phones are providing opportunities to learn in small bits of time as well as in dedicated sessions.
- Mind, body, and soul: Breakthroughs in brain research and cognitive science are being incorporated into learning. Evidence is mounting that physical fitness, happiness, diet,
overall wellness, and mindfulness (e.g., Goldie Hawn’s MindUp program) are associated with successful learning.

- Equity is the greatest concern and pervades all of the trends. For example, education technology has costs which limit access. While mobile technology is available to many, ten percent of students do not have smartphones.

**Discussion**

Thanos Patelis (HumRRO) facilitated a deeper discussion among panelists about common themes and the discussion questions.

*Personalized learning.* Content can be tailored to student preparation, interest, and ability. Learning will feel more purposeful, connected, and relevant. Fewer students will be seated in rows in classrooms on a rigid schedule. In high school, students may enroll in work training programs or participate in micro-internships. Teachers will serve as mentors. There is a need to change the traditional school organization/culture and provide teachers with the knowledge and skills to educate students in a new environment.

*Contextual data.* Is a student goal-focused or not? Using data about students’ goals can improve instruction. Contextual data (e.g., goals, interests, self-confidence) may provide clues as to why a student might be struggling and may also provide insights to inform how to individualize instruction.

*Equity.* Opportunity to learn pervades multiple areas. Cost and availability can be barriers to access educational technology and higher education.

*Big data.* Educational technology generates a lot of data. Educators need to learn how to analyze and use the data, taking a data systems point of view. Also, there is a need to teach teachers how to capture and document performance data on what students are doing in the classroom and how to use those data to improve classroom instruction and activities.

*Data dashboards.* Data dashboards can connect data from different sources, interpret multiple data points, and provide evidence of what students can do (versus cannot do).

*Micro-credentials.* Micro-credentials can be used by students and teachers. Students could earn a micro-credential when mastering a concept. Teachers can use their students’ micro-credentials to identify the skills acquired and those that need to be taught or re-taught.

*Competency assessments.* Students would benefit from measures of job-related skills to show their potential and demonstrate performance capabilities, particularly if the measures do not correlate to student background. Employers benefit because they have evidence of a job candidate’s skills. Educators can use competency data to mentor students on achieving goals.
Panelist Recommendations

As a wrap-up exercise, Thanos Patelis asked each panelist to make one recommendation for the Governing Board to consider.

Randy Bennett – Use NAEP’s national probability sample to describe what instruction is like at different levels for different types of students (e.g., students with disabilities, socioeconomic status) across time.

Karen Cator – Work toward a more coherent assessment system across NAEP and states.

David Conley – Endorse the work of the Ad Hoc Committee with a longer-term vision for NAEP to be bold in creating better items and measuring traditional content with greater precision.

Alana Dunagan – Develop innovative methods to measure flexibility, problem solving, and non-traditional skills that people will need in the future.

Devon Fidler – Look at partnering with prestigious organizations within the learning space that function outside of formal assessment, such as skunk works and incubators.

Nancy Lue – Use NAEP to assess the technology gap and equity issue in technology use outside of the classroom.

Reflections

Terry Mazany expressed his appreciation for the panelists’ insights. He noted that each expert presented similar ideas through a different lens; while this might have seemed repetitive, it actually reinforced the conclusions. The panelists convinced him that traditional education enterprise is collapsing in slow motion. Innovation outside of education is occurring at an accelerating pace. Learning might occur in smaller units such as micro-credentials.

Mr. Mazany discussed the high cost of traditional higher education and the trillion-dollar impact of student debt on the economy. He acknowledged the existence of prestige-based signaling that maintains inequity in the system. These are complex and challenging social issues. NAEP may be able to be a market signal by Governing Board priorities regarding what to measure and report on. He opined that perhaps NAEP can reinforce that prestige alone is not the gold standard.
Appendix A: Meeting Agenda, Attendees, and Panelist Biographies

Futurist Expert Panel
Thursday, June 21, 2018 1:00 pm – 4:00 pm PT
Room: Cypress A  * Hyatt Regency San Francisco Airport
1333 Bayshore Highway  * Burlingame, California, USA, 94010

Agenda

1:00 – 1:15 pm  Welcome, Introductions, and Overview of the Ad Hoc Committee
Terry Mazany, Chair of the Ad Hoc Committee on Measures of Postsecondary Preparedness
Overview of the Agenda and Goals for the Meeting
Thanos Patelis, HumRRO

1:15 – 2:45 pm  Panelist Perspectives and Initial Thoughts Regarding the Discussion Questions
A series of ten-minute presentations, each followed by a five-minute Q&A.
1:15 – 1:30     Randy Bennett (Educational Testing Service)
1:30 – 1:45     Karen Cator (Digital Promise)
1:45 – 2:00     David Conley (EdImagine)
2:00 – 2:15     Alana Dunagan (Clayton Christensen Institute)
2:15 – 2:30     Devin Fidler (Rethinkery Labs)
2:30 – 2:45     Nancy Lue (Advanced Education Research & Development Fund)

Questions for Discussion:
1. What are the trends you see that will define the future of learning and schooling?
2. What are the trends you see that will define the future of work and the skills that will be most valued by employers of the future?
3. What are the most promising technologies that will redefine education?
4. What things are most likely to disrupt how we think about teaching and learning?
5. What are the trends that most concern you, and why?

2:45 – 3:45 pm  Panel Discussion
Facilitated by Thanos Patelis

3:45 – 4:00 pm  Final Reflections
Terry Mazany

Conducted in Support of the National Assessment Governing Board’s Ad Hoc Committee on Measures of Postsecondary Preparedness
Attendees

Expert Panelists:
- Randy Bennett, Norman G. Frederickson Chair in Assessment Innovation in the Research & Development Divisions, Educational Testing Service
- Karen Cator, President and CEO of Digital Promise
- David Conley, President, EdImagine
- Alana Dunagan, Researcher for Higher Education, Clayton Christensen Institute
- Devin Fidler, Founder, Rethinkery Labs
- Nancy Lue, Co-Lead, Advanced Education Research & Development Fund

Governing Board Members:
- James Geringer, former Governor of Wyoming
- Carol Jago, Associate Director, California Reading and Literature Project at UCLA
- Terry Mazany, Chair, Ad Hoc Committee on Measures of Postsecondary Preparedness
- Dale Nowlin, Teacher and Mathematics Department Chair, Bartholomew Consolidated School Corporation, Columbus, Indiana
- Linda Rosen, former Chief Executive Officer, Change the Equation, Washington, DC
- Chasidy White, Director of Strategic Initiatives, Office of the Superintendent, Montgomery, Alabama

Governing Board Staff Members:
- Michelle Blair, Assistant Director for Assessment Development
- Bill Bushaw, Executive Director
- Lisa Stooksberry, Deputy Executive Director
- Lily Clark, Assistant Director for Policy & Research

HumRRO Staff Members:
- Monica Gribben, Senior Staff Scientist
- Sunny Becker, Principal Staff Scientist
- Thanos Patelis, Principal Scientist
Ad Hoc Committee Meeting on Postsecondary Preparedness

Panelist Biographies

Randy E. Bennett Ph.D. ............................................................................................................49
Karen Cator ..............................................................................................................................50
David Conley, Ph.D. ..................................................................................................................51
Alana Dunagan .........................................................................................................................52
Devin Fidler ...............................................................................................................................53
Nancy Lue ................................................................................................................................. 54
Randy E. Bennett Ph.D.
Norman O. Frederiksen Chair in Assessment Innovation in the Research & Development Division Educational Testing Service

Randy E. Bennett is Norman O. Frederiksen Chair in Assessment Innovation in the Research & Development Division at Educational Testing Service in Princeton, New Jersey. Bennett's work has focused on integrating advances in cognitive science, technology, and educational measurement to create approaches to assessment that have positive impact on teaching and learning. From 1999 through 2005, he directed the NAEP Technology Based Assessment project, which included the first administration of computer-based performance assessments with nationally representative samples of school students, and the first use of "clickstream," or log file, data in such samples to measure the processes used in problem solving. From 2007 to 2016, he directed an integrated research initiative titled, Cognitively-Based Assessment of, for, and as Learning (CBAL), which focused on creating theory-based summative and formative assessment intended to model good teaching and learning practice. Randy Bennett is president of the International Association for Educational Assessment (IAEA) (2016-), an organization primarily constituted of governmental and non-governmental nonprofit measurement organizations throughout the world, and immediate past president of the National Council on Measurement in Education (NCME) (2017-2018), whose members are individuals employed primarily in universities, testing organizations, state education departments, and school districts. He is a Fellow of the American Educational Research Association.
Karen Cator is President and CEO of Digital Promise and a leading voice for transforming American education through technology, innovation and research. From 2009-2013, Karen was Director of the Office of Educational Technology at the U.S. Department of Education, where she led the development of the 2010 National Education Technology Plan and focused the Office’s efforts on teacher and leader support. Prior to joining the department, Cator directed Apple’s leadership and advocacy efforts in education. In this role, she focused on the intersection of education policy and research, emerging technologies, and the reality faced by teachers, students and administrators. She began her education career in Alaska as a teacher, ultimately leading technology planning and implementation. She also served as Special Assistant for Telecommunications for the Governor of Alaska. Cator holds a master’s in school administration from the University of Oregon and received the 2014 College of Education Distinguished Alumni award. The American Association of Publishers has awarded Cator with the 2014 Visionary Award. She received her bachelor’s in early childhood education from Springfield College and received the 2015 Distinguished Alumna award. She is an Aspen Pahara Fellow, the past chair for the Partnership for 21st Century Skills and has served on boards including the Software & Information Industry Association-Education.
David Conley, Ph.D.
President, EdImagine
Professor of Educational Policy and Leadership in the College of Education at the University of Oregon
Director, Center for Educational Policy Research

David Conley is Professor of Educational Policy and Leadership in the College of Education at the University of Oregon where he directs the Center for Educational Policy Research. He is the founder and president of EdImagine, an educational strategy consulting company. Additionally, he founded and served for 12 years as CEO of the Educational Policy Improvement Center, EPIC (now Inflexion). He recently completed an appointment as Senior Fellow for Deeper Learning under the sponsorship of the Hewlett Foundation.

Dr. Conley is a national thought leader in the areas of college and career readiness, student ownership of learning, systems of assessment, educational accountability, and the future of education and the economy. He has published multiple articles and policy briefs as well as three books in these areas. His most current book, published by Harvard Education Press, is entitled The Promise and Practice of Next Generation Assessment.

He serves on numerous boards and advisory committees including as a member of the technical advisory committee of the Smarter Balanced Assessment Consortium (SBAC) and the Illinois State Board of Education Accountability Technical Advisory Committee, and as a founding board member of New Meridian, which now manages the PARCC assessments. Additionally, he chairs the New Meridian Steering Committee. Previously, he co-chaired the Validation Committee for the Common Core State Standards.

He has conducted multiple major research studies for the Association of American Universities, the College Board and its Advanced Placement program, the International Baccalaureate, and the National Assessment of Governing Board. He has most recently studied next generation systems of assessment, new indicators of college readiness, and new methods to determine career readiness.

Before entering higher education at the University of Oregon in 1989, Dr. Conley spent 20 years in the public-school system in a variety of roles including teacher and co-director of two alternative schools, a site and central-office administrator, and an executive in a state education agency. He is a first-generation college attendee who received his AA from Cabrillo College, his BA from the University of California, Berkeley, and his MA and PhD from the University of Colorado, Boulder. He grew up on the central coast of California, where he spent a great deal of time at the beach.
Alana Dunagan
Researcher, Higher Education, Clayton Christensen Institute

Alana leads the Institute’s higher education research and works to find solutions for a more affordable system that better serves both students and employers. In this role, Alana analyzes disruptive forces changing the higher education landscape. Her research includes studying business model innovations, public policies, and investment strategies that can give rise to new and sustainable postsecondary models.

Prior to joining the Christensen Institute, Alana spent ten years in institutional investment management working on behalf of nonprofits, particularly colleges and universities. She worked as an investment consultant for Slocum, and spent five years with Macalester College managing their $700 million endowment. She holds a BA in Economics and Political Science from Macalester College and an MBA from the Harvard Business School.
Devin Fidler
Founder, Rethinkery Labs

Devin has worked with senior leaders at dozens of Fortune 1000 companies to systematically explore emerging issues and technologies, and to analyze their potential impacts. His ongoing work at Rethinkery Labs, including developing tools for “self-driving” management, has been covered by HBR, the New York Times, Wired and a number of other publications. He argues that today, companies themselves are a technology on the verge of disruption. Prior to founding Rethinkery, Devin founded and led the Future of Work and Future of Learning programs at the Palo Alto-based Institute for the Future.

Devin is a frequent speaker at gatherings of business leaders and others interested in the transformation of work and organizations. He approaches projects from a strongly international perspective, having lived and worked in several countries throughout his career.
Nancy Lue  
Co-Lead, Advanced Education Research & Development Fund

Nancy Poon Lue is currently co-leading the exploration of a national Advanced Education Research & Development Fund on behalf of the Chan Zuckerberg Initiative and the Bill & Melinda Gates Foundation. She is also a Partner and Secretary of the Board of Directors of the venture philanthropy organization Silicon Valley Social Venture Fund (SV2). Previously, she served as Executive Director at the venture capital firm Global Silicon Valley (GSV) and was the inaugural General Manager of the EdTech Lab at GSVlabs. During the Obama Administration, Nancy was a Senior Advisor at the U.S. Department of Education where she led the development of the agency’s five-year strategic plan. Nancy is a Senior Fellow with the American Leadership Forum-Silicon Valley and sits on the Advisory Board of the AT&T Aspire Accelerator and the GreenLight Fund-Bay Area. She earned her B.A. and Ed.M. from Harvard College and Harvard Graduate School of Education.
Appendix B: Panelist Presentations

Note: These slides are the intellectual property of the presenters and should not be used or distributed for purposes beyond this Committee without permission.
Trends in the Future of Learning

- Be technology based, making greater use of complex tasks, games, simulations
- Based on more modern underlying models of cognition and learning
- Be personalized in terms of:
  - Accessibility
  - Competency level
  - Background and interest
  - Learning goal
- Include (or give greater emphasis to) "new" competencies, e.g.:
  - Self-regulated learning
  - Citizenship, civic engagement
  - Crisis-natural competence
  - Using technology tools for problem-solving
- Include traditional competencies:
  - Knowledge acquisition and construction
  - More (not less) important for technology
- Include focus on cross-cutting skills within the disciplines
  - Communication, critical-thinking
- Embed assessment within instruction, including automated analysis and feedback

Thoughts on the Future of Education and Work

Randy Bennett
Educational Testing Service
Princeton, NJ 08541
rbennett@ets.org

Presentation as a member of the Futurist Expert Panel at the meeting of the National Assessment Governing Board's Ad Hoc Committee on Measures of Postsecondary Preparedness, San Francisco, CA, June 2018.

Trends in the Future of Work

- Continued automation of many types of manual, cognitive, and social-interactional work
- Pervasiveness of technological tools for problem solving as:
  - Aids requiring constant proximal human interaction (Excel)
  - Extensions allowing for remote reach (drones)
  - Assistants: Carry out this subtask (Siri)
  - "Intelligent" implementers: Work independently with human QC

Overview

1. What are the trends you see that will define the future of learning and schooling?
2. What are the trends you see that will define the future of work and the skills that will be most valued by employers of the future?
3. What are the most promising technologies that will redefine education?
4. What things are most likely to disrupt how we think about teaching and learning?
5. What are the trends that most concern you, and why?
#### Most Likely Challenges to Disruptions for Teaching and Learning

- In K-12, the challenges are still greater than the disruptors
  - Level of, and extent of variation in, quality of teaching
  - Level of, and extent of variation in, school technology
  - Variation in funding for education by locale
  - Grade-based organization of schooling
  - Local control
  - Little coherence, massive inefficiency due to lack of chance for economy of scale
  - Size (3rd largest country in the world)
  - Concerns for privacy of student data
  - Concerns over the corporatization of education
  - Public indifference, even antipathy, toward rigorous expectations and toward addressing inequality

#### Skills Most Valued by Employers

- Using technology tools for problem solving—i.e., to create value by being able to use:
  - Aids requiring constant proximal human interaction (Excel)
  - Extensions allowing for remote reach (drones)
  - Assistants: Carry out this subtask (Sin)
  - “Intelligent” implementers: Work independently with human QO
  - Being able to, individually and in collaboration with others, locate, evaluate, integrate, synthesize, apply, and construct knowledge (i.e., to learn)
  - Being able to communicate, educate, and help others make effective decisions

#### Most Likely Disruptors

- People
  - What factors will make educators, policy makers, parents, students, and public advocate for, and accept, change?

#### Most Promising Technologies for Education

- Technologies that increase opportunities for remote social interaction
  - Learning is a social activity
- Adaptive learning (intelligent tutoring) combined with human instruction
- Simulations, games, virtual reality that pose tasks and situations similar to the ones students must learn to negotiate as proficient practitioners in a domain
- Analytics to help adapt instruction, guide students in managing their learning, help teachers improve instruction and its management
Trends of Most Concern

- Personalization
  - Equity: differential foci of instruction by demographic group
- Idea of replacing end-of-unit assessment with embedded formative assessment
- Use of AI (without sufficient human oversight) for consequential decision-making purposes
  - Where explanation is important, current approaches to AI are insufficient for making decisions that affect life chances
  - EU GDPR requires provision of an explanation
1. What are the trends you see that will define the future of learning and schooling?

- Personalization
- Learner Variability (advancements in learning sciences)
- Competency based learning (+ performance assessment)
- World Challenges (e.g., UN SDGs)
- Workforce Changes

Global Goals
What are the trends you see that will define the future of work and the skills that will be most valued by employers of the future?

**Artificial Intelligence**

- Ability to learn
- Work with others
- Flexibility and comfort with complexity
- Creativity and solution development
- Compassion and Thinking

What are the most promising technologies that will redefine education?

**Augment Human Performance**

- Data and instrumentation (like location & weather) - Adaptive
- Improved diagnostics and embedded assessment
- Moving from (average and comparison) to precision and accuracy
- Virtual and augmented reality
- Open Education Resources - organized, findable and contextualized
What things are most likely to disrupt how we think about teaching and learning?

- Learning Sciences
- Improvement Science
- Research and Evidence
- Advanced R&D - Pasteur’s Quadrant

What we know - Learning Sciences

- Learning is Developmental
- Challenging & Achievable
- Total Environment
- Practice Practice Practice
- Personal & Meaningful
- Habits of Mind
- Prior Knowledge
- Emotion Matters
- Social Interaction
- Unique Jagged Profiles

Learner Variability

What are the trends that most concern you, and why?

- Lack of respect for and disenfranchised teachers
- Under-resourced schools
- Cost of higher education
- Assessments that fall far short of the full picture
- Issues of inequity
- Digital Learning Gap (Access - Participation - Powerful Use)
Conley Presentation
NAGB Presentation

Linda O. Dean
Senior Advisor for Future
National Association of Governing Boards

the questions

1. What are the trends you see that will define the future of learning and schooling?
2. What are the trends you see that will define the future of work and the skills that will be most valued by employers in the future?
3. What are the most promising technologies that will redefine education?
4. What things are most likely to change our thinking about teaching and learning?
5. What are the issues that most concern you, and why?

Trends that will define the future of learning and schooling?

- Personalization of access to knowledge
- Universality of the learning process
- Importance of instructors as mentors and peer-to-peer
- Emphasis on success in learning
- Social learning
- Complex profiles of learners
- Importance of adaptability over fixed skill set
- Need to be able to learn in complex and uncertain environments
- Adaptability to learning

Trends that will define the future of work

- Gray area
- Low skill vs. high skill
- Fast immerse, fast move
- Strategic vs. legacy businesses
- Globalization vs. localization
- Struggle vs. flexibility
- Standards vs. product
- Innovation vs. standardization vs. customization/complexation
skills that will be most valued by employees of the future

- self-direction
- data analysis and interpretation
- ability to identify and implement innovative, alternative
  solutions to uncertain, multi-stakeholder problems
- communication
- teamwork
- problem formulation
- problem solving
- ability to work with a wide variety of people
- attention to detail, having precise procedures
- ability to read technical manuals, interpret graphs, charts, diagrams
- all types of computer skill including web-based skills

the most promising technologies that will redefine education

- instantiation
- understanding the physiology of learning
- eye tracking
- biofeedback
- fluidity
- gamification
- simulations, virtual reality
- virtual environment
- artificial intelligence
- self-learning, self-awareness
- self-learning, self-awareness
- criticism of learning
- self-learning, analysis, personalization

things are most likely to disrupt how we think about teaching and learning

- self-directed learning
- emphasis on basic skills in our current educational system
- human interactions and adaptation of knowledge
- simulation, serious games
- self-directed learning
- distance education
- personalization of learning
- self-learning, feed back, awareness
- equity issues
- decreasing economic and educational policy limits
- growing gap in measures of "perceived knowledge" at all levels
- widely different definitions of what "prepared" means
- varying influence between economic and social groups
- ever-increasing pace of disruptive economic change
- lag between changing economic and social structure and education's ability to adapt
- and the role of formal and informal social networks that underlie social mobility functions and social memory.
Disruption could create major benefits for students and society:

- Radically Affordable
- Accessible to All
- Knowledge + Skills
- Workforce Aligned
- Modular + Scalable
- Public + Private
Today's Landscape: Enrollment

Skills matter more than ever
- Prepare workers for the jobs they don't know exist
- Skills required today differ from those growing today
- What are those skills?

But our economy is more dependent on skilled workers than ever before

Lots of new, innovative players emerging
1) Online learning is quickly becoming ubiquitous in higher education.

- Fall 2016: 30.8% of undergrads and 34.6% of graduate students are learning online.
- Online learning is especially prevalent among those who otherwise wouldn't be able to access higher education.
- About half of all students enrolled in school were learning online.

Access and equity, affordability

2) Learning in smaller denominations than the degree...

...this creates big opportunities.

College is still the currency of the labor market...

...but, as Ryan Craig says, it's a currency system with $10,000 bills and no smaller denominations.

Automation and technology are changing the nature of work—and that change will be continuous.

Learning won't end with college graduation.
3) Workforce alignment built into program and curriculum design

- Building relevant certificates into programs
- Using industry experts (rather than academic experts) to design curriculum
- Creating explicit learn-to-work and learn-to-network opportunities
- Experiential learning
Fidler Presentation
Biotech
Nanotech
Greentech

**Much Bigger: Organizational Technologies**

---

**Preparing for Next Generation Organizational Technologies**

Devin Fletter | Devin@futurethinkerslabs.com
There Are Many Ways to Organize

- 4000 years ago: First legal structures for companies
- 900 years ago: Trade guilds encode industry across Europe
- 250 years ago: Industrial revolution and modern companies emerge
- 190 years ago: Assembly lines, globalization etc.
There Are Many Ways to Organize

- 1000 years ago: First legal structures for companies
- 1300 years ago: Industrialization and modern companies emerge
- 200 years ago: Multinationals and globalization
- 130 years ago: Assembly lines, globalization etc.

Digital Organization is Special
Human Task Routing

Use emerging technologies to:
activate, deactivate & reconfigure
resources
where they are needed &
when they are needed

The Frontiers of “OrgTech”

A Collective Intelligence Engine

© Work that looks for people,
instead of people looking
for work
© Using platforms to find the
best matches available
© Collective Intelligence
Surfacing insights
Training, National Interest, and the New OS

- Orchestration becomes a more powerful skill
- Many more established industries can expect competitors built along these lines - Transition and national interest

Thank You!
Devin Rober | Devin@RethinkeryLabs.com

RETHINKERY LABS
National Assessment Governing Board Futurist Panel
June 21, 2018
Nancy Poon Lue

1. Return on Education (ROE)
6. Mobile

7. Mind, Body, Soul
Summary of the Expert Panel Meeting with State Education Officials
June 28, 2018

National Assessment Governing Board
Ad Hoc Committee on Measures of Postsecondary Preparedness

To support the charge of the Ad Hoc Committee on Measures of Postsecondary Preparedness, the Human Resources Research Organization (HumRRO) convened a small volunteer panel of education officials responsible for their state’s assessment and/or accountability. The meeting was conducted in partnership with the National Assessment Governing Board (Governing Board) and the Council of Chief State School Officers (CCSSO). The focus group was conducted on June 28, 2018, in San Diego, California during the CCSSO-sponsored National Conference on Student Assessment. The purpose of the focus group was to gather information about states’ definitions of postsecondary preparedness/readiness and their efforts to develop and use indicators of postsecondary preparedness/readiness.

The focus group participants included Chris Janzer, Michigan; Russell Keglovits, Nevada; Shelley Loving-Ryder, Virginia; Vaughn Rhudy, West Virginia; Michael Sibley, Alabama; Jenny Singh, California; Allison Timberlake, Georgia; and Vince Verges, Florida. Ms. Loving-Ryder and Mr. Sibley participated in the panel as both state experts and members of the State Policy Task Force, which is jointly convened by the Governing Board and CCSSO.

In attendance were Governing Board members Tyler Cramer and Joseph Willhoft; Governing Board staff members Michelle Blair, Lily Clark, Sharyn Rosenberg, and Lisa Stooksberry; CCSSO staff members Fen Chou and Scott Norton; and HumRRO staff members Sunny Becker, Monica Gribben, Thanos Patelis, Sheila Schultz, and Arthur Thacker.

An overview of the Governing Board and the charge of the Ad Hoc Committee on Measures of Postsecondary Preparedness, along with the agenda and the logistical information for the meeting, were sent to the participants as read-ahead materials. The meeting agenda is at Appendix A.

Thanos Patelis, HumRRO Principal Scientist, started the meeting by reviewing the agenda and goals. Lily Clark, Governing Board Assistant Director for Policy and Research, welcomed everyone and provided an overview of the Governing Board’s Strategic Vision initiative to “develop new approaches to measure the complex skills required for transition to postsecondary education and career,” which led to the creation of the Ad Hoc Committee on Measures of Postsecondary Preparedness and the impetus for this focus group meeting.

Mr. Patelis facilitated a discussion among the participants that highlighted the following guiding questions:

- How does your state define college and career readiness?
- Did your state consult with industry groups to define career readiness?
- What measure(s) does your state use to assess career readiness?
- Is military service a component of postsecondary readiness in your state?
- How does your state use noncognitive measures?
• Are there innovative or nontraditional indicators that your state might use to measure or report on students’ college and/or career readiness (e.g., student interest, micro-credentials earned, work-based learning)?
• What NAEP reporting on postsecondary readiness would be useful to states?

Following is a general summary of the information provided by this group of state assessment and accountability experts on definitions, activities, and indicators of postsecondary preparedness/readiness.

Definitions
The state officials offered examples of definitions of college and career preparedness/readiness used in their respective states. It was evident from the examples that states have a variety of definitions for college and career readiness. The definitions and indicators for college readiness were separate from those of career readiness. Most of the definitions for career readiness explicitly included “soft skills,” such as communication, collaboration, problem solving, and business practices. The state officials acknowledged the importance of soft skills to college and career readiness while also noting the challenge they pose in developing and measuring indicators related to these skills.

The definitions of college and career preparedness/readiness represented by the participating state officials varied in certain aspects and included the following:

• Two states defined college readiness as students who enroll and succeed in college courses without remediation.
• The use of benchmarks on college entrance and placement tests serve as a default definition of college readiness.
• Career readiness can be defined as obtaining a job that pays a living wage, which varies by location.
• Career readiness in several states was defined by a set of credentials from a career and technical education (CTE) program that did not include inter- and intra-personal skills. However, some other states included soft skills, such as inter- and intra-personal skills and business skills, in their definitions.
  o In one state, the inclusion of service learning was part of the secondary school experience that contributed to a career ready diploma seal.
  o In another state, career readiness was defined as acquiring specific skills from CTE programs as well as successful performance on assessments that represented specific skills (e.g., National Occupational Competency Testing Institute) and experience in a simulated workplace program.
• One state described the development of college and career readiness standards that defined specifically what is meant by college attendance and students’ understanding of the available career fields.
• Military readiness was offered as a postsecondary option that involves a set of cognitive and physical requirements, which is seen as an indicator of readiness in some state accountability plans.

A couple of state officials commented how they would welcome a definition of college and career readiness from the Governing Board.
Learning Opportunities and Interventions
Several state officials described the following efforts for students to acquire college and career readiness skills:

- States work with schools and industry to develop diplomas to certify technical career skills.
  - The diploma is earned through CTE programs, work-based learning, industry/credential exams, or portfolios.
  - One state developed career ready diploma seals that reflect cooperation between CTE programs and industry to introduce service learning and experiences for students to acquire industry-specific technical and broad interpersonal and intra-personal skills (e.g., leadership, collaboration, communication skills).

- Programs to prepare students for career readiness are designed to take advantage of local industry and involve the cooperation and input of businesses likely to hire postsecondary students.

- Schools encourage or adopt dual enrollment initiatives to increase student access to college-level courses and experiences.

- Soft skills, such as communication and leadership skills, are taught through service learning, student organizations, work-based learning, and simulated work environments.

- One state’s goal is to prepare students for college or a career by ensuring they are agile in facing an environment where the requirements are not always known.

- One state official indicated that the state department of education is (and should be) flexible in facilitating local education agencies to develop pathways for students that are relevant for local conditions and situations.
  - As an example, one school district described a multi-national company that moved into the municipality with plans to add an international business pathway for students. Students who complete designated international business courses and activities earn a career ready seal on their diplomas.

Data and Indicators
The state officials identified sets of skills important for college and career readiness. Some commented on the difficulty in measuring certain skills from both practical/logistical and technical/measurement perspectives. One state official opined that it is easier to measure college readiness than career readiness. Many state officials noted the difficulty with career readiness data is twofold: (a) the skills to be assessed are multi-faceted in nature and (b) there are practical limitations in identifying measurable indicators for all facets.

The skills explicitly mentioned, especially for career readiness, include business practices, collaboration, leadership, communication, creative problem solving, argument and reasoning, designing solutions, time management, and intellectual curiosity.

Several state officials indicated the Governing Board could contribute to the measurement of the soft skills important for indicating career readiness, particularly if provided at the state level. One official, however, encouraged the measurement of both college and career skills, but also cautioned that one consequence of reporting these skills by state is how industry may use them to target or avoid certain states for opening corporate and business locations.

State officials offered various comments and suggestions about data related to college and career readiness:
• Geographic differences reported in relevant career skills were based on the types of local industry and available jobs. States want data at a regional level.
• Some soft skills are not easily defined or measured (e.g., time management, intellectual curiosity).
• Student level data on absences, credits, and required course attainment can serve as proxies for some soft skills.
• A portfolio of artifacts (in the form of certificates, work-based learning, etc.) or experiences (advanced courses, dual credit) can be used as an indicator of college and career readiness.
• A concern about equity in terms of (a) opportunities to learn and (b) distribution of funds to offer college and career readiness opportunities (test fees) was expressed.
• Student service learning could be used as a relevant data point.
• One suggestion was for states to support and incorporate local accountability plans and metrics that involve school-specific indicators of important constructs such school culture, climate, and other environmental measures.
  o Examples of using school climate and school culture surveys were reported.
• Indicators used in state accountability plans included attendance, course participation, college entrance and placement test scores, and certification test results.

Various comments were offered about the measurement of college and career readiness:

• College readiness is easier to measure than career readiness.
• Soft skills typically are not included in state standards, so what to measure becomes a challenge.
• Measures should be general (versus specific) to remain relevant over time.
• Soft skills should be measured early (e.g., age appropriate elementary and middle school skills) to allow time for students to close gaps and attain common school and workplace skills. Early measurement would provide schools with data to monitor student learning and acquisition of these important life skills.
• States would like to see best practices in providing, documenting, and measuring college and career readiness skills.
  o For example, is there evidence that students who earn certificates are successful?
• A couple of state officials commented that the Governing Board is in a unique position to develop a measure(s) of soft skills at the state/national level.
• It would be a tremendous contribution if the Governing Board created a single definition inclusive of both college and career readiness as well as developed indicators to measure those skills.
Appendix A: Meeting Agenda and Attendees
Discussion of State Efforts on College and Career Readiness

Thursday, June 28, 2018, 7:30 – 8:50 AM PST
Room: Cobalt 520 (Level 5)
Hilton San Diego Bayfront
San Diego, California

Agenda

Purpose: Identify and discuss states’ current and innovative practices regarding college and career readiness to inform the National Assessment Governing Board’s effort to “Develop new approaches to measure the complex skills required for transition to postsecondary education and career.”

7:30 – 7:45 AM Breakfast & Introductions

7:45 – 8:00 AM Overview of the National Assessment Governing Board’s Initiative on Postsecondary Preparedness
Lily Clark, Assistant Director for Policy and Research
National Assessment Governing Board

8:00 – 8:50 AM Discussion of State Efforts on College and Career Readiness
Thanos Patelis, Facilitator, HumRRO

Guiding Questions:

- How does your state define college and career readiness?
- Did your state consult with industry groups to define career readiness?
- What measures does your state use to assess career readiness?
- Is military service a component of postsecondary readiness in your state?
- How does your state use non-cognitive measures?
- Are there innovative or non-traditional indicators that your state might use to measure or report on students’ college and/or career readiness (e.g., student interest, micro-credentials earned, work-based learning)?
- What NAEP reporting on postsecondary readiness would be useful to states?

8:50 AM Thank you and Adjourn
Attendees

State Officials (Department of Education)
Chris Janzer, Michigan
Russell Keglovits, Nevada
Shelley Loving-Ryder, Virginia
Vaughn Rhudy, West Virginia
Michael Sibley, Alabama
Jenny Singh, California
Allison Timberlake, Georgia
Vince Verges, Florida

CCSSO Staff Members
Fen Chou
Scott Norton

National Assessment Governing Board Members
Tyler Cramer
Joe Willhoft

National Assessment Governing Board Staff Members
Michelle Blair
Lily Clark
Sharyn Rosenberg
Lisa Stooksberry

HumRRO Staff Members
Sunny Becker
Monica Gribben
Thanos Patelis
Sheila Schultz
Arthur Thacker
The summary of the Young Adult Expert Panel meeting is forthcoming.
Work of the Future – 2030

Overview of Jobs of the Future
History shows major changes in the occupational landscape and pace of life with each of the four industrial revolutions (Vale, 2016). The first industrial revolution, characterized by the steam engine, led to greater dispersal of jobs as those requiring machine power were not restricted to locations with wind or flowing water to power mills. Electricity and mass production brought about the second industrial revolution, leading to a surge in manufacturing jobs and supporting industries such as transportation, sales, and business. The advent of the digital age, the third industrial revolution, gave us the ability to collect and process massive amounts of data quickly and opened up new jobs related to computers and technology innovation. Now, we are entering the fourth industrial revolution, highlighted by the internet of things and artificial intelligence (Choi, 2017; Vale, 2016).

Throughout history, the introduction of new technologies has led to changes in jobs, from replacing workers with machines to changing how people perform their job to creating new occupations. According to futurists, this trend will continue. Policy analysts predict up to 47 percent of jobs in the United States could be automated between 2017–2037 (Bakhshi, Downing, Osborne, & Schneider, 2017; Houser, 2017). Opportunities will become limited in many industries, mostly in low- or medium-skill jobs, as automation reduces the number of humans needed to perform routine tasks. Further, business leaders and strategists predict that 50 percent of the occupations of 2014 will no longer exist in 2025 (Andrew, Ip, & Worthington, 2014). Technology, automation, artificial intelligence, and other innovations that have yet to be developed will lead to new occupations and jobs.

Atkinson and Wu (2017) take a different perspective of technological disruption, suggesting that others have based their doomsday predictions of rampant job loss on “faulty logic and erroneous empirical analysis.” Instead, they calculate, from 2010 to 2015, approximately six technology-related jobs were created for every 10 lost, the lowest share of jobs lost to technology of any period since 1950 to 1960.

While there will likely be changes in jobs and occupations of the future, what those changes will be, the extent and pace of changes, and the impact on employees entering or currently in the workforce are equivocal. In this report, we review the research related to potential changes in the workplace and highlight forecasts of jobs of the future.

Projections of Shifts in Jobs
Prediction of widespread unemployment due to technological advances is nothing new. For example, in the 1930’s, John Maynard Keynes predicted large-scale job loss associated with new technologies (as cited in Frey & Osborne, 2013). Recently in the United States, automation has been replacing jobs faster than it can create them (Atkinson & Wu, 2017; Brynjolfsson & McAfee, 2011), although perhaps not as quickly as some suggest. Autor, Levy, and Murnane (2003) found that as industries use automated technology to reduce the cost of performing routine cognitive and manual tasks, they hire more people to perform nonroutine cognitive tasks.

The occupations in which people are or will be employed are expected to shift, but this does not necessarily mean current jobs will be totally eliminated. As Manyika (2017a) reports, at least 30 percent of activities for most occupations could be automated using current technology. Assuming in many current occupations certain activities or tasks will be automated, current jobs
will change and more people will need to work with technology. Although some employees may lose their jobs because automation will drastically eliminate the need for human skills, integration of technology will help other workers perform their job better or enable them to be more efficient or productive. For still other workers, the demand for their skills may increase or the nature of what they do and how they accomplish tasks at work will change.

**O*NET Projections**

O*NET OnLine (National Center for O*NET Development, 2018) is a rich source of “detailed descriptions of the world of work.” There is a wealth of data available to those looking for work or interested in changing careers, as well as support for workforce development and human resources professionals, researchers, and policy analysts.

Using 2016–2026 employment projections from the Bureau of Labor Statistics, O*NET includes a set of Bright Outlook occupations. Twelve of the Bright Outlook occupation categories (including 20 distinct occupations) are expected to grow rapidly with an employment increase of 10% or more and are forecasted to have 100,000 or more job openings between 2016 and 2026 (see Table 1). O*NET identifies occupations linked to the green economy, focused on reducing environmental risks and initiating sustainable development without degrading the environment. Green jobs identified in O*NET are those where changes are expected in job demand, including work requirements such as tasks performed or worker qualifications such as knowledge, skills, and credentials needed for employment in these positions.

**Table 1. O*NET Bright Outlook Occupations with Rapid Growth and Numerous Job Openings**

<table>
<thead>
<tr>
<th>Major Occupation Group</th>
<th>Occupation Category</th>
<th>Green Economy Sector</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business and Financial Operations</td>
<td>Accountants and Auditors</td>
<td></td>
</tr>
<tr>
<td>Education, Training, and Library</td>
<td>Teachers and Instructors, All Other (includes Tutors)</td>
<td></td>
</tr>
<tr>
<td>Healthcare Practitioners</td>
<td>Registered Nurses (includes Acute Care Nurses, Advanced Practice Psychiatric Nurses, Critical Care Nurses, and Clinical Nurse Specialists)</td>
<td></td>
</tr>
<tr>
<td>Healthcare Support</td>
<td>Home Health Aides</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nursing Assistants</td>
<td></td>
</tr>
<tr>
<td>Personal Care and Service</td>
<td>Personal Care Aides</td>
<td></td>
</tr>
<tr>
<td>Food Preparation and Serving</td>
<td>Cooks, Restaurant</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Combined Food Preparation and Servicing Workers, Including Fast Food</td>
<td></td>
</tr>
<tr>
<td>Building and Grounds Cleaning and Maintenance</td>
<td>Janitors and Cleaners, Except Maids and Housekeeping Cleaners</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Landscaping and Groundskeeping Workers</td>
<td></td>
</tr>
<tr>
<td>Sales</td>
<td>Sales Representatives, Services, All Other (includes Energy Brokers)</td>
<td>X</td>
</tr>
<tr>
<td>Construction and Extraction</td>
<td>Construction Laborers</td>
<td>X</td>
</tr>
</tbody>
</table>

Source: O*NET OnLine
Several researchers have mined the O*NET data to make predictions about the future of jobs, identifying ones expected to increase and ones to decrease in the future. Bakhshi et al. (2017) used O*NET’s importance ratings in foresight exercises to generate input for a machine learning model, with the goal of mapping O*NET knowledge, skills, and abilities variables to future occupational demands. In the United States, the model predicts increased demand for teachers from pre-school through high school and post-secondary. Animal care workers, legal professionals, and engineers round out the top five occupations with expected increased demand.

**Jobs Expected to be Lost or to Decrease**

Job loss and decrease due to technology is evident all around us. Grocery stores offer multiple lanes where customers scan their own purchases, monitored by a single cashier. Only a few years ago there were multiple cashiers serving the customers. Financial institutions offer more and more automated functions such that their customers need to interface with a person less often than ever before. For example, customers can now use a mobile application to deposit a physical check without leaving home. Andrew, Ip, and Worthington (2014) predict customer work will disappear and many middle management positions will no longer exist in 2025. Frey and Osborne (2013) describe greater use of data and algorithms to computerize cognitive tasks such as fraud detection, health care diagnostics, legal document review, and financial advice. With automation expanding into more cognitively-advanced occupations, demand for individuals with certain professional skills, such as financial analysts and law clerks, are predicted to decline.

Frey and Osborne (2013) used O*NET data to study how susceptible jobs are to computerization. Using data for 702 occupations, they modeled the potential for jobs to be automated within 10–20 years. Their model predicted workers are most likely to be replaced with technology in occupations that involve transportation and logistics, office and administrative support, manufacturing, and service.

Bakhshi et al. (2017) did not use their model to predict decreased demand, but rather to predict low probability of increased demand. Those in skilled and semi-skilled trades, such as woodworkers, printing workers, metal and plastic workers, and other production occupations, were at the top of the future low demand list. Financial clerks received low ratings comparable to those in the trades.

**Jobs Expected to Increase**

Bakhshi et al. (2017) expect growth in professional occupations that require creative, digital, design, and engineering expertise. In addition to creativity, Osborne and Frey (n.d.) suggest growth in jobs that require social intelligence and manipulation, such as iOS and Android developers, social media interns, big data architects, data scientists, user interface/user experience (UI/UX) designers, Zumba instructors, and beachbody coaches. Further, strong interest in environmental sustainability is expected to benefit individuals employed in architectural and green occupations (Bakhshi et al.). Also, they foresee increased roles for people specializing in work reorganization, such as management analysts and training specialists.

**New Jobs to be Created**

Jobs requiring creative intelligence and social and emotional intelligence are predicted to be added to the economy, as are positions requiring the ability to leverage artificial intelligence (AI; Andrew, Ip, & Worthington, 2014). New jobs using creative or social and emotional intelligence or AI are expected to be more fulfilling than current jobs.
Generally, specific details about future jobs are scarce. Wagner (2011) discusses 70 jobs likely to exist in 2030. These jobs will be created through (a) retrofitting or adding new skills to existing jobs, (b) blending or combining functions from different jobs or industries, and (c) problem solving or creating new jobs to solve a problem. Types of jobs that might be added through retrofitting could support commercial space travel, such as space construction, space suit repair, space junk recyclers, astro-teachers, and exozooologists. By blending careers, the future might include environmental health nursing to treat patients exposed to toxins. To provide authoritative news in an era when anyone can publish online may lead to authority-journalists who specialize in an occupation and are cross-trained to report about their field. To solve future problems, we may hire digital footprint managers or digital archaeologists or future-guides. Wagner mentions occupations in the sustainability and green energy industries may be added, such as green career coach, autonomous vehicle operator, energy harvester, drone dispatcher, smart car interior designer, smart road designer/engineer, and smart road sensor control monitor/analyst. Gordon (2011) predicts there will be new careers inspired by nanotechnology, such as bio-botic physicians and bio-botist assistants to integrate biological functionalities and implanted nano-robotics to extend life.

**Drivers of Change**

The literature discusses three major drivers of projected shifts in jobs of the future—technology (Frey & Osborne, 2013), artificial intelligence (Manyika, 2017b), and social changes (Bakhshi et al., 2017; Manyika, 2017a). With changes in jobs come adjustments in the workplace. Experts predict that workplace culture and processes will shift as well as career paths and how people learn the necessary job skills needed to perform jobs of the future.

**Impact of Technology**

Literature is replete with observations of the accelerating impacts of technology in recent decades, including predictions this acceleration will continue. Baby boomers remember a world when communication required a phone call that was timed when both parties were available to speak or a letter that took days to be delivered; researching a topic involved going to the library or referencing a home copy of encyclopedia volumes; and getting a flat tire meant a hike to find a pay phone. Today’s young people are digital natives. They cannot imagine a world before e-mail allowed asynchronous communication; the internet offered a wealth of instantaneous information at one’s fingertips; and cell phones connected individuals to worlds beyond measure. The explosion of technology is expanding in multiple directions—and quickly.

Bakhshi et al. (2017) employed an innovative approach to predicting job trends by first paneling experts in “foresight workshops” and then inputting their expert judgments into a machine learning system. Their analysis of the experts’ judgments identified three key trends in technological change. First, fears about the impact of automation on employment are enduring. Second, estimates of the impact of future automation range from 9–47 percent of U.S. employment. Third, technology can amplify human performance and bring about new occupations and sectors.

For its 21st annual survey of CEOs worldwide, PwC interviewed 1,293 CEOs in 85 countries, including 104 from the United States, in October and November 2017 (Ryan, Sapin, Rao, & Ampil, 2018).

Based on these interviews, U.S. CEOs were hiring for broadly relevant digital skills and collaborative, creative, and efficient work styles. About two-thirds (63%) of those who were hiring found it more difficult to identify qualified workers than before. Responses to this same
survey indicated that artificial intelligence (AI) will be the innovation of the next two decades. CEOs predict that many workers will need AI literacy.

One of the challenges of a rapidly changing work environment is the ability of workers to keep pace. While new entrants into the workforce will grow up and attend school immersed in state-of-the-art technology, continued innovation ensures even these digital natives—those who have an advantage over older, digital immigrants who completed school before digital technology became omnipresent—will require ongoing training to stay current on technological knowledge, awareness, and skills. Employees who joined the job market prior to many of the current technological advances are already challenged with staying up to date. Two approaches to keeping tenured employees abreast of the latest technological developments are upskilling and reskilling.

**Upskilling**

When an employee upskills, that individual learns new skills to improve performance on the job or to adapt to new requirements of the job. Upskilling has the advantage of retaining experienced employees, a positive outcome as these employees are a known commodity to the employer, absent the risks of employing a new hire who may not be a good fit. Knowledge of corporate procedures, norms, and expectations eliminate the need for orientation and start-up time, and reduce the probability of missteps. Retaining seasoned employees also supports the maintenance of institutional memory, which can be crucial as an organization evolves and grows.

Training to upskill employees may be sought by the employee, imposed by the employer, or both. The PwC’s Workplace of the Future study found three-quarters of respondents expressed willingness to update their own skills. At the same time, most responding CEOs acknowledged an ongoing responsibility to upskill their employees (Ryan et al., 2018).

PwC’s Annual Global CEO Survey specifically investigated the employer’s perspective on upskilling. Nearly two-fifths (39%) of respondents reported initiating or using continuous learning initiatives to provide development paths for employees to gain skills.

**Reskilling**

When an occupation becomes obsolete or the changing nature of the position no longer suits an employee, reskilling may be in order. More disruptive than upskilling, reskilling is training an employee to perform an entirely different job.

Results from PwC’s annual survey of CEOs indicate companies that “reinvent their own talent” by reskilling their employees will have an edge by creating pathways for employees to better contribute to data-driven initiatives, which may lower costs and improve the customer experience among other impacts (Ryan et al., 2018). However, the U.S. lags other large economies (e.g., Germany, China, Japan) in assuming responsibility for retraining after automation (i.e., robots and AI) has been introduced into a job. The authors conclude that, while automation will result in job losses, over time those will be generally offset by new jobs. They note that “retraining workers to work with the support of AI will be important to future economic success” (Ryan et al., 2018, p. 15)

**Working with Data**

Alec Ross, author of Industries of the Future, provides an historical perspective of the workplace. He describes land as the raw materials of the agricultural age, followed by iron in the industrial age, and data in the information age. He posits that whoever owns, controls, and/or
can harvest meaning from data will define the future workplace. Ross (2016) emphasizes the sheer quantity of data being produced in recent history and the opportunity for data analytics to mine those data. For example, he notes that “90 percent of the world’s digital data has been generated over the last two years” (page 154). He opines that the sum of “all data from paintings on cave walls through 2003, we now produce every two days” (Ross, 2017).

PwC issued a report on the workforce of the future, using findings from a survey of 10,000 people in China, India, Germany, the United Kingdom, and the United States on how they think the workplace will evolve. From the survey findings, the authors developed four “Worlds of Work” for 2030 to describe hypothetical future scenarios defined along two continuums: collectivism and fragmentation. Authors concluded the increasing use of digital platforms and AI mean data are key. With augmented intelligence, humans and machines collaborate to make decisions. Uniquely human traits of emotional intelligence, creativity, persuasion, and innovation become more valuable. Adaptability will become increasingly important as work changes (PwC, 2017).

**Human-Technology Interactions**
Not only are data produced and stored at astounding rates, but individual access to such data through technology is expanding. Ross (2017) asserted that 20 billion networked devices were in circulation in 2017. He projected this number will reach 45 billion in 2020. This growth will likely not produce a steady expansion across all markets, but rather result in bursts of growth in traditional areas that have not been as impacted by the digital economy, such as transportation or mining.

PwC’s Annual Global CEO Survey predicts that businesses will initiate upskilling initiatives to teach employees the skills they need to augment their own work with the support of technology. The authors contend that companies will infuse AI into all aspects of their business, not just technology-related areas (Ryan et al., 2018).

**New Technology Jobs**
Technology jobs such as software engineers are on the rise, but two other trends may result in new technology jobs. First, the blending of AI technology with a human component, or augmented intelligence, may open opportunities for technology-enhanced versions of jobs that are available today (PwC, 2017).

Second, Ross (2016) points out an increasingly popular conviction that the opportunities of the future will no longer rigidly distinguish technical fields from liberal arts or humanities. He suggests hybrid studies will become more prevalent, such as a combination historian/electrical engineer or political scientist/computer scientist. He describes the thinking of Toomas Ives, President of Estonia: “…domains previously occupied only by people with backgrounds in the liberal arts, like government, will become increasingly occupied by people with more background knowledge in science and technology” (page 246).

**Impact of Artificial Intelligence**

**Types of Artificial Intelligence**
Although people may mean different things when they refer to artificial intelligence (AI), they generally mean the use of computers to perform tasks that require cognition and learning without programming the steps of the task. Often, AI is used to refer to machine learning, “where computers are taught or self learn how to recognize things” (Shaw, 2017). Bughin et al. (2017) describe other types of AI, including computer vision, autonomous vehicles, natural language, smart robotics, and virtual agents.
Machine learning is intensive, for the humans who must provide the “training data” and for the computer to process the information. Shaw notes that machine learning has many applications, such as predicting nefarious behavior or mechanical breakdown and identifying possible disease in 3D radiology images. Research is underway to explore the use of AI to make machine learning more efficient and accurate.

Shaw expects computer vision, using cameras to infer what they are seeing, to become the most prevalent type of sensor. Computer vision will be integral for self-driving cars and other autonomous vehicles such as self-driving trucks, buses, trains, and ships. Autonomous flying drones, which may be used for package delivery or to aid in aerial search and rescue, also will benefit from computer vision.

Natural language processors are familiar to many as they ask Siri for directions or to settle a debate. Smart home devices such as Google Home or Alexa are natural language processors. Once these devices understand what a person has said or written, Shaw states that a virtual agent is the next step. The virtual agent can help the human, provide financial advice, perform basic health diagnosis, or guide an individual through steps of an activity or job. Smart robotics are in use today, especially in manufacturing. Shaw expects robotics to become more prevalent in medicine, cleaning, stocking, agriculture, and food service in the future.

Machine learning is but one way of many to categorize AI. Hintze (2016) defines AI using a hierarchy from type I-reactive machines (e.g., Deep Blue, IBM’s chess supercomputer) to type II-limited memory (e.g., self-driving cars monitor information over time) to type III-theory of mind (e.g., understanding that thoughts and emotions affect behavior) to type IV-self-awareness (e.g., being aware of oneself).

**New Artificial Intelligence Jobs**

Research continues to advance AI (Bughin et al., 2017; Hintze, 2016; Shaw, 2017), with the implication that jobs developing and studying AI will continue to grow as the technology is incorporated into more daily life routines. Research firm Gartner, as cited in Singh (2017), predicts by 2020 more jobs will be created by the expansion of AI than will be lost. They estimate, that although AI will be responsible for the loss of 1.8 million jobs between 2018 and 2020, AI will create 2.3 million jobs. Healthcare, the public sector, and education will lead the way in incorporating AI into their sectors.

Increasing use of digital platforms and AI mean data will be key to creating new AI jobs (PwC, 2017). With augmented intelligence, humans and machines must collaborate to make decisions. Singh (2017) expects one in five workers will rely on AI to assist them in their jobs by 2022. It will be important for people to learn to work with and alongside AI machines.

**Impact of Social Changes**

**Globalization**

Globalization refers to the increasing interconnectedness of the world, both economically and politically. Along with automation, globalization is viewed as one of the main factors shaping the future workforce (Bernstein, 2016; Simon, 2016).

Companies operating on an international scale may have financial incentive to move jobs from the U.S. to other countries. This has been demonstrated historically through the loss of low-skilled manufacturing jobs due to offshoring (Hatzichronoglou, 2005). Today, higher skilled jobs also run the risk of offshoring, including computer-oriented science, technology, engineering, and mathematics (STEM) jobs (Lim, 2016).
Technological advances are closely linked with globalization’s impacts on the workforce. “Labor-linking” technology allows geographically dispersed people to vie for the same job, creating a competitive environment that could potentially drive wages down (Basu, 2016). However, there remain numerous higher-skilled jobs that are less subject to this threat, particularly those in healthcare and service industries that require face-to-face interactions (Blinder, 2007).

The potential for interaction with clients, customers, and coworkers from across the globe has implications for the skills that are valuable as well as valued. For example, employees may find it easier to negotiate the workplace when they have the skills needed to communicate effectively with geographically dispersed people from a range of sociocultural backgrounds. Employers are increasingly recognizing the value of cultural competence and communication skills among new hires (Vozza, 2016), especially when those skills are needed to perform future jobs that involve interaction on a global scale.

**Environmental Sustainability**

Environmental policies have long been linked to the reduction of jobs in specific industries (e.g., coal), though a causal link is up for debate (Morgenstern, Pizer, and Shih, 2001). The availability of jobs in such industries in the U.S. may in fact be limited by a decreased demand for fossil fuels that has resulted from advances in energy efficiency technology (Magill, 2017). Regardless of the mechanism at work, there is reason to believe the jobs of the future will continue to be shaped by both policy and consumer behavior related to environmental sustainability.

Beyond contributing to the obsolescence of some jobs, the focus on environmental sustainability continues to create new jobs and to change jobs that currently exist. Job opportunities for innovation related to environmental sustainability may increase as individuals and corporations alike seek to reduce energy consumption and waste (Bakhshi, Downing, Osborne, & Schneider, 2017). Companies that are changing practices to reduce their environmental footprint may create jobs for “sustainability professionals” who will take on the role of managing company resources (Hamilton, 2012). The National Center for O*NET Development has identified green economic sectors, green increased demand occupations, green enhanced skills occupations, and green new and emerging (N&E) occupations, many of which will likely boast increased job opportunities in coming years. Green enhanced skill jobs are those in the existing occupation that require significant changes due to the impact of the increased focus on environmental sustainability (O*NET, 2018).

**Demographic and Population Patterns**

As of 2016, foreign-born workers constituted nearly 17% of the U.S. labor force (Bureau of Labor Statistics, 2017). By 2060, approximately 20% of the total national population is expected to be foreign-born (Colby & Ortmann, 2015). Some raise concerns about the potential loss of jobs by American-born workers to immigrants (Hoban, 2017). Others argue the rising immigrant population will increase opportunities for U.S.-born workers, as immigrants frequently perform low-skilled jobs that are complementary to, and increase the productivity of, work performed by other Americans (Greenstone & Looney, 2012). However, many immigrants also hold advanced degrees, particularly in STEM fields (Solis, 2011), and could therefore play a crucial role in meeting the demand for highly skilled workers.

As working Baby Boomers draw closer to retirement age, there is concern over the loss of the knowledge and skills of the overall labor pool (Burke & Ng, 2006). Globally, the ratio of non-working age people to working age people appears to be on the rise (Bakhshi, Downing, Osborne, & Schneider, 2017). This trend may be counterbalanced by policy changes that raise
retirement ages or provide incentives for older workers to remain on the job (Lerman & Schmidt, 1999).

The overall aging of the population has implications for available jobs. Jobs in healthcare and the production of goods and services targeting the needs of older citizens are on the rise (Singh, 2015). However, the influx of highly-educated Millennials into the workforce has its own implications. These workers are anticipated to bring a new set of expectations of their employers, including demands for improved working conditions and human resources policies (National Academies of Sciences, Engineering, and Medicine, 2017). At the same time, there is rising concern regarding this full subpopulation’s preparedness with the skills required in the ever-changing world of work (KRC Research, 2014).

**Education and Training**

It has been estimated that most children entering primary schools today will work in job types and roles that don’t yet exist and that will be characterized by the need for not only technological, but also social and analytical skills (World Economic Forum, 2016). It is anticipated workers of the future will hold an increasing number of jobs over their lifetime (Pompa, 2015). These factors, coupled with increasingly rapid technological change, will necessitate a continuous process of education and training throughout these future workers’ careers (Karoly & Panis, 2004). This suggests the need for consideration of both the education and training offered to students prior to their entry into the paid labor force, as well as how systems for continued education and training will be implemented and sustained.

Numerous innovative approaches to preparing students with the in-demand middle level skills needed to perform jobs of the future are expanding in their implementation. Career and technical education (CTE) programs, apprenticeships, early college high schools, and career academies are among the approaches that seek to bolster the skill levels of Americans entering the workforce for the first time (Joint Economic Committee Democrats, 2018). On-the-job training (OJT) models are another innovative approach that provides incentives to employers to hire lower-skilled workers and offer them targeted training while they engage in paid labor, as well as offer continued training to allow for career advancement (Kobes, 2013).

**Equity Issues**

Many anticipate the trends of globalization and automation will lead to increasing inequality, as wages for highly skilled workers rise while low- and unskilled workers will compete with both automation and workers located in other countries (The Foundation for Young Australians, 2017). Other areas of concern regarding equity in the workplace relate to gender and age. Women who seek to both parent and work continue to face potential wage reductions and loss of skill development when they take time off for family leave (O’Marah, 2018). Some anticipate women will be disproportionately impacted by job losses due to automation (Hayasaki, 2017). Aging workers may face threats to their continued employment over issues related to healthcare costs and age-related disabilities (National Bureau of Economic Research, 2018).

**New Social-Oriented Jobs**

Futurists envision new positions will be created to do work that has never been done before. Most of these jobs will develop in response to shifts in the marketplace or they will be created because of advancing technologies. Envisioned jobs include those that harness the power of social media to create tailored experiences for customers or clients (Wagner, 2010). With more companies using social media to connect with customers and to expand their presence in the market, employees’ experience with and understanding of social media will be increasingly valued by employers (Kumar, Bezwada, Rishika, Janakiraman, & Kannan, 2016). Companies
and organizations will need to monitor, maintain, and improve their online presence, and new positions will likely be created for that purpose (University of Kent, 2018).

**Workplace of the Future**

When examining the workplace of the future, Frey and Osborne (2013) convened human experts in machine learning to classify a subset of jobs according to the likelihood of their “automatability.” Through analysis of O*NET variables as proxies for three irreplaceable attributes (i.e., perception and manipulation, creative intelligence, and social intelligence) they developed a model to predict the automatability of the full set of O*NET occupations. Results indicated that 47 percent of U.S. employment can be classified as high risk for automation within the next decade or so.

The Guardian’s Workplace Benefits Study (2017) defines four top trends impacting the workforce in 2018 and beyond. Each of these trends is related to technology:

- Technology is enabling an on-demand workforce;
- Automation is requiring an enhancement of workforce skillsets;
- Employers are reinventing talent recruitment; and
- Varying workplace demographics require different strategies for adoption.

**Workplace Culture**

**Agile Workforce**

As organizations are required to respond quickly to changes in an increasingly globalized and technologically advanced world, they seek an agile workforce that is similarly capable of responding to unanticipated change with speed and flexibility (Breu, Hemingway, Strathern & Bridger, 2001). Workers of the future may be expected to rotate among a variety of roles and tasks, as employers seek to find the skills needed for a specific task at a particular time (Wadors, 2018). As companies leverage a variety of work models (e.g., ad hoc teams, crowdsourcing, independent contractors) to meet their needs, workers may find themselves entering into many different types of nontraditional work arrangements (Green, 2014). Andrew, Ip, and Worthington (2014) expect an increase in distributed work places. Continual reskilling will be a key element in sustaining high levels of agility (Lyons, Blitz, & Whittall, 2017).

**Less Structure and Predictability**

Careers have been traditionally viewed as a progression of jobs, often upward through a predictable, hierarchical structure (Lyons, Schweitzer, & Ng, 2014). Careers of the future will likely unfold in less hierarchically structured environments, where there will be increased interconnectedness among departments and where individuals may assume different job roles depending on the context of the work at hand (Heerwagen, 2016). Job tasks themselves are expected to be less structured and predictable as new technologies replace once rote and predictable duties with ones that require abstract thinking and flexibility (National Academies of Sciences, Engineering, and Medicine, 2017).

**Sharing Economy**

More and more, modern day consumers and workers engage in short-term economic transactions around services that involve sharing some material good (e.g., car, living space) or skill for monetary compensation (Sundararajan, 2016). Also referred to as the gig economy, platform economy, access economy, or collaborative consumption, this sharing economy is
anticipated to increase exponentially over the coming decades (Yaraghi & Ravi, 2016). Such work arrangements have both potential positive and negative consequences for workers of the future. It can be argued that individuals will be empowered by the sharing economy to go into business for themselves and gain returns on their assets. On the other hand, the sharing economy removes protections that workers have enjoyed under more traditional work arrangements (Lamberton & Rose, 2012).

**Continuous Learning**

McKinsey & Company (2017) recommend that workers of the future be prepared to be lifelong learners. McKinsey Global Institute (MGI) partner Susan Lund explained, “For young people today, what’s clear is that they’re going to need to continue to learn throughout their lifetime. The idea that you get an education when you’re young and then you stop and you go and work for 40 or 50 years with that educational training and that’s it—that’s over. All of us are going to have to continue to adapt, get new skills, and possibly go back for different types of training and credentials. What's very clear is that what our kids need to do is learn how to learn and become very flexible and adaptable.”

Guardian (2017) recommends that employers address the need for continuous learning through experiential, retraining, and cross-training programs, as well as mentoring, e-learning opportunities, and tuition assistance.

Ross (2016) opines the U.S. adoption of free education until the age of 18 was appropriate as long as a high school graduate could get a job in a “port, factory, mine or mill—a middle class job.” However, in the information age, he suggests we know the pace of change demands that we be lifelong learners.

**Flexible and Non-Traditional Career Paths**

Predictions regarding future career paths are wide ranging. Popular “wisdom” has long asserted that younger generations no longer expect to join an employer after high school or college and stay with that same employer until retirement. Lyons, Schweitzer & Ng (2015) analyzed the career mobility patterns of four generations and found that job mobility increased with each successive generation. Specifically, “The magnitude of the differences was large, as Millennials [born 1980 or later] had almost twice as many job and organizational moves per year as the generation Xers [1965-1979], almost three times as many as the Boomers [1946-1964], and 4.5 times as many as the Matures [born prior to 1946]” (page 16). However, this change in job mobility does not reflect an increase in employee turnover from one employer to another, but rather increased movement through various positions within a company. They postulate that technology, among other factors, may make some positions obsolete. The authors conclude the traditional career model is still strong and the “oft-cited truisms about the ‘new’ or ‘modern’ careers may be exaggerated” (page 18).

Intuit & Emergent Research (2017) predict that by 2021, 9.2 million American workers will derive at least some of their income as independent contractors operating within a “gig economy”—situations in which organizations establish short-term contracts on an as-needed basis. This is a substantial growth projection relative to the 3.9 million in 2016. McKinsey Global Institute (2016) estimates that 20–30 percent of individuals of working age in the U.S. and the European Union conduct independent work.

This trend is facilitated by technology that allows a job incumbent to be geographically distant from the employer; the advantages to an organization of selecting the best candidates for a given project, without a long-term commitment; and the ability to increase and decrease staff
levels as demand warrants. This is further enabled by current and planned features in job-employee matching software such as Monster.com, Aftercollege.com, and Taskrabbit and networking sites such as LinkedIn (Brynjolfsson & McAfee, 2016).

Interdisciplinary Teams
Based on research by Burkus (2016), some organizations encourage employees to engage in more face-to-face communication in an effort to increase problem solving and decision making efficiency (as cited in Colbert, Yee, & George, 2016). As a result, workplaces evolve to provide more flexible space for collaborating and working in teams (Giang, 2015). Experts from Unum Limited’s Futures100 network (2014) foresee more conversation and debate, either face-to-face or on conversation-based platforms. Employees will need to blend skills and disciplines when working with others. They will collaborate with each other rather than compete. Workers will need listening skills and to display empathy, and build relationships to enable collaborative and interdisciplinary ventures.

Summary of Themes of Work and Workplace of the Future
When it comes to work of the future, change is the only certainty. However, this review of relevant literature points to some overarching themes that provide a solid base for making predictions about the world of work that today’s kindergartners will need in 2030 when they graduate from high school. This world will likely look very different from the world of work their parents were prepared for, both in terms of the available jobs and the work environment in which those jobs are carried out.

Jobs of the future will undoubtedly involve technology. From searching job openings, to performing job tasks, to receiving professional development, interacting with new and emerging technologies will be a distinctive feature of future jobs. Fields that had previously been quite separate may be blended in new ways, and existing jobs may be blended with new technologies to create positions we’ve never seen (think: space junk recyclers!).

The high school graduates of 2030 will set out on a career pathway characterized by change. Whether they work independently through the gig economy, or move among multiple employers or across multiple departments or projects, workers of the future will likely find themselves part of an increasingly diverse and dispersed workforce. Jobs will be continually evolving to meet changing demands and to incorporate the latest innovations. Ongoing training will be a necessary component of future jobs. Employees will need to adapt and embrace life-long learning to be successful in the workplace.

With some sense of what the future holds for work and the workplace, it becomes clear expected changes in jobs and job environments will correspond to changes in associated skills. A key next step to ensuring that students graduate high school in 2030 prepared for the next step on their postsecondary pathway is to identify the skills and abilities needed to successfully perform the jobs of the future.
**References**


Appendix I. Literature Review: Work of the Future


Appendix J. Literature Review: Skills of the Future
Skills of the Future – 2030

Overview of Skills of the Future
Students completing kindergarten in 2018 will graduate from high school in 2030. As the work and workplace of the future change, so will some of the skills students need for success in postsecondary activities. In Work of the Future – 2030, Gribben, Becker, and Dickinson (2018) described research related to potential changes in the workplace and forecasts of jobs of the future. Although we do not know with certainty what types of jobs will be available in the future, researchers and business analysts use trends to predict the types of jobs and skills that they expect high school graduates will need for employment in the future (e.g., Grover, 2018; Lara, 2018; McKinsey & Company, 2017; P21, 2016).

This review follows that of the Work of the Future – 2030 and answers a range of questions about the landscape of postsecondary skills in 2030. What skills will students need following graduation from high school in 2030? Are there skills common across jobs? What skills do secondary students need either for matriculation into college or entering the workforce?

Skills high school graduates in 2030 will need to succeed along postsecondary pathways must correspond to the jobs of the future. Technology is expected to play a large role in future jobs. With the advent of the digital age, there is a recent emphasis on skills required to use and interact with new devices and applications. Employers will need programmers and innovators to develop new technologies to tackle more difficult challenges and improve efficiency and cost effectiveness. High school graduates of 2030 will likely find themselves part of an increasingly diverse and dispersed workforce. As the workplace and postsecondary institutions become more distributed and global, employees and students will need facility with collaboration tools as well as socio-emotional skills for working with diverse colleagues. Jobs will continually evolve to meet changing demands and to incorporate innovations. Employees will need to adapt and embrace life-long learning. Automation will replace some jobs and disrupt certain industries; future jobs are likely to require complex problem solving and troubleshooting which are not easily automated. Skills that enable individuals to work with and use technology, communicate with others, and continually adapt and learn will be necessary for the high school class of 2030.

In this literature review of the skills of the future, we provide a brief introduction to college and career preparedness. We present a structure for organizing the discussion of skills of the future followed by research on each of the skills. Similar to the Work of the Future – 2030 literature review, we include projections of shifts in future skills. We discuss skills expected to increase in demand through 2030 and new skills expected to be added. The report concludes with a summary of themes of skills for the future.

Integrated Framework for Postsecondary Preparedness
There has been much discussion and work looking at college and career preparedness either by bifurcating college and career preparedness or by assuming them to be the same (ACTE, 2010; Conley, 2011; National Center for O*NET Development, 2018; P21, 2016). In addition, global statements frequently are made about preparedness without considering context. Patelis (2018) has proposed an integration of popular career and college preparedness frameworks to address the limitations in other organizing schemes to cover both career and college.

Focusing on the skills at the intersection of career and college preparedness, we use a unified framework to represent postsecondary preparedness. This offers a way of thinking about the overlapping nature of the skills needed for career and college in a more integrated manner.
rather than as separate or redundant constructs. The National Research Council (NRC, 2011; Pelligrino & Hilton, 2012) organized postsecondary skills into three categories: cognitive, intrapersonal, and interpersonal. These clusters encompass knowledge and skills\(^{41}\) needed for life after high school. To address the more complex and multidimensional skills of the future, however, we have added another category – blended skill sets.

As the NRC defined them, cognitive skills for the 21st century involve (a) cognitive processes and strategies, (b) knowledge, and (c) creativity. Cognitive processes and strategies encompass critical thinking, complex problem solving, and analysis and interpretation. Knowledge covers academic areas such as reading, writing, and science, technology, engineering, and mathematics (STEM). With the ubiquity of digital devices, skill in using or developing digital tools is an important addition to the set of cognitive skills of the future. Creativity includes innovation and creative skill sets.

Interpersonal skills, sometimes called social skills, require complex communication, teamwork, and collaboration. These range from communication and collaborative problem solving to cooperation and perspective taking. With expected increases in diverse and globally dispersed workplaces of the future, we have added cultural awareness and sensitivity to the interpersonal skills needed in the future.

Intrapersonal skills cover intellectual openness, and work ethic and conscientiousness. Intellectual openness includes adaptability, personal responsibility, and continuous learning. Work ethic and conscientiousness includes initiative, productivity, and professionalism.

The following sections look at the future of (a) cognitive skills, (b) interpersonal skills, (c) intrapersonal skills, and (d) blended skill sets. Table 1 presents a list of the skills within each of the categories. The skills discussed here are not intended to be comprehensive. Rather, we have focused the discussion on skills that figure most prominently in recent thinking of the future of work.

\(^{41}\) Knowledge refers to what a person knows and understands. Skill refers to what a person can do.
Table 1. Postsecondary Preparedness Skills for 2030

<table>
<thead>
<tr>
<th>Cognitive Skills</th>
<th>Interpersonal Skills</th>
<th>Intrapersonal Skills</th>
<th>Blended Skill Sets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foundational Academics</td>
<td>Communication (including Listening, Conversation, and Persuasion)</td>
<td>Time Management</td>
<td>Learning Agility</td>
</tr>
<tr>
<td>STEM</td>
<td>Relationship Building</td>
<td>Efficiency</td>
<td>New Media</td>
</tr>
<tr>
<td>Critical Thinking</td>
<td>Cultural Sensitivity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complex Problem Solving</td>
<td>Understanding Other People’s Perspectives</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Creativity</td>
<td>Collaborative Problem Solving</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Innovation</td>
<td>Social and Emotional Intelligence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digital Tools</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistical Literacy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computational Thinking</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Cognitive Skills**

**Foundational Academics**
Basic skills in education, namely literacy and numeracy, are well defined, taught, and measured. They are foundational to the acquisition of knowledge and skills (Peterson et al., 2001) and the performance of most tasks (e.g., Durak & Saritepeci, 2018). Economic research shows that improving basic skill proficiency has a dramatic effect on important societal outcomes, such as wage growth (McIntosh & Vignoles, 2001) and social development (Hanushek & Woessmann, 2008), further demonstrating the criticality of these skills to important outcomes.

As work evolves over the coming decades, literacy and numeracy will remain important, although the way they are used may change. For example, tasks requiring basic literacy and numeracy, such as data entry and cashiering, will likely become automated, thereby reducing the demand for these skills by as much as 25% from 2016 to 2030 (Bughin et al., 2018). As work becomes less structured and predictable and more team- and project-based, literacy will remain critical for effective acquisition and communication of information. For example, the demand for more advanced literacy and numeracy skills, such as are used in analytic and communications activities, is expected to increase by almost 10% (Bughin et al., 2018).

Other academic areas beyond literacy and numeracy include mathematics and science. These have become increasingly important and are often grouped together with technology and engineering to form their own category of Science, Technology, Engineering, and Mathematics (STEM).

**STEM**
Skills in the STEM areas are frequently included in discussions of the future of work. While great effort is currently being put into projecting and developing STEM skills for the future, the concept of STEM as an educational discipline dates to at least two National Science Foundation reports.
in the 1990s (Advisory Committee to the National Science Foundation Directorate for Education and Human Resources, 1996; 1998). Since that time, researchers worldwide have attempted to define STEM skills and strategies for developing them in the current, and future, workforce.

While many of these discussions project shortages of STEM skills, other research suggests that there is both a shortage and a surplus of STEM skills, driven by occupational and geographic mismatches between individuals’ skills and specific job openings. STEM is a very broad concept and not all skills within it are equal in demand. Xue and Larson (2015) noted that, due to the heterogeneous nature of STEM, some skills are in short supply. These include very rare skill sets (e.g., Ph.D.-level nuclear engineering), as well as more common (e.g., data science, software development) and overlooked STEM skills (e.g., skilled trades). Xue and Larson also identified surpluses of STEM skills, such as biomedical and chemical engineering, as well as geographical variation in supply and demand (e.g., software engineering skills are in high demand in California).

Other researchers point out that basic STEM literacy will be critical even for non-STEM jobs. In fact, some definitions of basic skills include STEM literacy as a basic skill or assume that most job candidates will have at least basic proficiency with STEM skills (Cunningham & Villaseñor, 2016; Roberts & Bybee, 2014; Zeidler, 2014) due to its perceived relevance for a world of technology-immersive work.

There is disagreement regarding the profile of skills defining STEM. Some definitions delineate specific STEM disciplines along with associated higher-order cognitive skills. Carnevale, Smith, and Melton (2011) and Jang (2015) empirically identified the skills associated with STEM occupations using the Occupational Information Network (O*NET\(^{42}\)) taxonomy. O*NET classifies its nearly 1,000 occupations into 5 job zones describing how much preparation (education) is required, job climate (e.g., working conditions, recognition, independence), and potential for future growth. Results of the STEM occupation crosswalk included a wide range of skills including both content knowledge (e.g., math, chemistry, and other scientific and engineering fields) and a broader set of cognitive skills relevant to the future of work (e.g., complex problem solving, deductive and inductive reasoning, mathematical reasoning, and facility with numbers).

Other definitions focus on the interdisciplinary components of STEM as they apply across disciplines and occupations:

STEM skills and knowledge are interdisciplinary in nature, being based on the integration of the formerly discrete disciplines of science, mathematics, engineering and technology. The aim of STEM skills is to enhance people’s competency in work and/or life and more generally respond to societal demands on technology.

STEM skills belong to the group of technical skills. They are a combination of the ability to produce scientific knowledge, supported by mathematical skills, in order to design and build (engineer) technological and scientific products or services. Although STEM skills overlap with basic and higher order cognitive skills, they merit separate treatment in a policy-oriented context in order to target specific requirements in the education and labor market. STEM skills and knowledge

\(^{42}\) O*NET (https://www.onetonline.org/) is developed under the sponsorship of the U.S. Department of Labor/Employment and Training Administration (USDOL/ETA).
cannot be directly measured by current discipline-specific classifications. 
(Siekmann & Korbel, 2016)

Siekmann’s (2016) “House of STEM” model integrates these perspectives. The “House of STEM” includes basic numeracy, literacy, and socioemotional (e.g., curiosity) skills as a foundation for the kinds of technical and higher-order cognitive skills identified by Carnevale et al. (2011).

In addition, other researchers advocate for broader conceptualizations of STEM that include skills from the arts and humanities (i.e., STEAM). These models reflect the importance of visual and verbal communication to both the discipline-specific technical skills and the higher-order cognitive skills comprising the House of STEM (Land, 2013). For example, modern approaches to data analytics rely heavily on data visualization, a skill that clearly draws upon the arts. The arts also help develop higher-order cognitive skills, such as critical thinking, creativity, perspective taking, and divergent thinking, that are critical to scientific advancement and innovation (Daugherty, 2013; Kahn & Zeidler, 2016). These researchers caution that focusing solely on a narrow definition of STEM skills may not adequately prepare us for the future.

**Critical Thinking**

Critical thinking is important for learning and for evaluating new information (Mumford, Peterson, & Childs, 1999), making it vital to performing work that does not follow prescribed rules (i.e., work that is not likely to be automated). In a future world of work characterized by expanding technology and artificial intelligence, humans’ unique capacity for critical thinking will set them apart from machine learning algorithms (Barnett, Lawless, Kim, & Vista, 2017; Pistrui, 2018). In fact, job recruiters across the major industries have identified critical thinking skills among the less common, but more desired skills among business school graduates (Levy & Cannon, 2016).

Glaser (1985) viewed critical thinking as a process involving careful consideration of a problem through the application of logical inquiry methods. McPeck (2016) identified the critical features of critical thinking, six of which distinguish critical thinking from other higher cognitive skills:

1. Involves the evaluation of statements/evidence and the methods used to derive them
2. Is limited in application to the context of a specific discipline, field, or problem; it cannot be learned or applied in the abstract
3. Does not require the thinker to reject established norms or ideas
4. Is informed by evidence but requires judgment
5. Includes active problem solving, not just evaluation
6. Is not redundant with logic or rationality.

**Complex Problem Solving**

In recent years, employers have consistently rated problem solving as the most essential competency for career preparedness (National Association of Colleges and Employers, 2017; Thompson, 2016). CEOs from across industries and around the globe expect work environments to continue to grow in terms of complexity (IBM, 2010), and complex problem solving is anticipated to replace physical abilities and core skills in a wide variety of jobs of the future (World Economic Forum, 2016). Thirty-six percent of all jobs are expected to require complex problem-solving as a core skill by 2020 (Thompson, 2016). The ability to address
complex problems is particularly important in work environments characterized by rapid change (Middleton, 2002).

Complex problems are those that involve multiple goals along with many possible courses of action, all of which may shift in an environment that is dynamic (Fischer, Greiff, & Funke, 2012). Complex problem solving includes not only cognitive, but also emotional and motivational elements. It is a dynamic process in which the pathway to the solution might be more informative about the problem solver than reaching the solution itself (Dorner & Funke, 2017).

**Creativity**

U.S. CEOs are looking for employees with creativity (Ryan, Sapin, Rao, & Ampil, 2018). Creativity is a skill often identified as important in a rapidly changing workplace, and a capability that machines do not possess. But the assumption that creativity is a purely human skill has been challenged. Early creativity assessments measured divergent thinking or resistance to functional fixedness. These assessments included tasks such as presenting an object and asking the examinee to generate as many uses of that object as possible, within a specified timeframe. Individuals who generated a larger number of unique uses, regardless of elegance or complexity, received higher scores. This task measured the ability to generate novel, or heretofore nonexistent, ideas.

One technique for creating novel ideas is to simply produce combinations of familiar ideas. For example, a common brainstorming exercise to generate ideas for creative writing is to create index cards—either electronic or physical—each with a noun or adjective, and then pick two or three cards at random to create an unusual combination to spark an original idea. Computers are certainly well suited to this sort of rapid, rote activity. However, while this approach demonstrates an ability to generate novel ideas it may result in nonsensical combinations.

Boden (2003) builds upon this definition of creativity as “the ability to come up with ideas or artifacts that are novel and valuable” (emphasis added). Similarly, Mumford (2003) describes creativity as “new and useful.” Under this framework, creativity requires the generation of unique ideas as well as the ability to evaluate those ideas. Evaluating creative ideas is a complex and domain-specific task that requires a wealth of information. Not only must this conceptual space include sufficient details about the field to evaluate whether ideas are sensible, but metrics are necessary to assess the value of an idea. These values are specific to content domains and may change over time.

Frey and Osborne (2013) point out that research literature reveals examples of software generating and evaluating creative solutions. For example, Harold Cohen produced a drawing program, AARON, in the mid-1970s in an attempt to answer the question, “What are the minimum conditions under which a set of marks functions as an image?” Over decades, AARON evolved into an extensive producer of visual art with an internal feedback system that has been displayed in art galleries (Cohen, 1995).

Similarly, David Cope produced the initial version of the Experiments in Music Intelligences (EMI) software program in the early 1980s to analyze the style of a given musical composer (e.g., Bach, Mozart, Prokofiev) and then generate an original composition in the same style (da Silva, 2003). This software required domain-specific knowledge of tone systems, phrase structure and length, rhythm, movement structure, etc. and the ability to evaluate the balance between unity and variety.
Despite these domain-specific examples of computer generated creative products, Frey and Osborne (2013) conclude that “it seems unlikely that occupations requiring a high degree of creative intelligence will be automated in the next decade” (p. 26).

Brynjolfsson and McAfee (2011) suggest that the combination of current economic trends, specifically job growth trends and wage stagnation at the median range, provides opportunities for creative entrepreneurs. While increased technology may eliminate some career paths, it also opens potential paths to create ways to use technology to support mid-skilled workers and add value.

**Innovation**

The Future Laboratory conducted a study to identify trends in workplaces, employers, and employees (Unum United, 2014). They interviewed a range of experts from its Futures100 network—including academics, authors, scientists, and social scientists—and also surveyed 1,000 employees for their perspectives on the trends identified via these interviews. Respondents indicated that future workplaces will be people-centric and will foster innovation. Employees will need to blend skills and disciplines.

Innovative employees can only thrive in an organization that accepts innovative ideas. While innovation has been touted for years as an important trend, recent studies call specifically for “responsible” or “disciplined” innovation. Sull (2015) appeals for the implementation of a simple set of rules to serve as constraints in the innovation process. He acknowledges that despite guardrails designed to manage innovation, some failures are still inevitable, but that incorporating discipline serves to increase efficiency and improve the odds of successful innovations. He characterizes an appropriate set of simple rules as (a) few in number, (b) applicable to a well-defined activity or decision (rather than a broad corporate principle that is too vague to be actionable), (c) tailored to the culture and norms of the organization, and (d) sufficiently flexible to allow creativity and discretion. He suggests that an organization might employ rules for various purposes, including to select innovations, define how to innovate, and help community members innovate together. He cites successful innovators ranging from Zumba Fitness to the Defense Advanced Research Projects Agency (DARPA).

A corporate reputation for innovative thinking also can attract new employees. Brown and Martin (2015) describe the success of Innova Schools, which brought affordable education to Peru. Innova drew upon the results of a stakeholder study to develop a technology-enabled education system that valued the “guide on the side” rather than the traditional “sage on stage” model of instruction. Invited, ongoing feedback from school leaders, teachers, and parents helped to introduce continuous improvements, some of which were fundamental shifts from the expected outcomes. This approach resulted in scalability and job growth. The authors note that “Because Innova had a reputation for innovation, teachers wanted to work there, even though it paid less than the public [school] system” (p. 9).

**Digital Tools**

While the construct of skill with digital tools is not well defined, there are two aspects of digital skills that are unique: the skills required to (a) create and (b) use digital tools. The creation of digital tools, such as artificial intelligence and machine learning, requires STEM, analytic, and computational thinking skills. More specific digital skills, such as web development, are likely to rapidly evolve with the advent of new technologies, methods, and languages.

Basic digital skills are useful for developing other skills. For example, some medical training now relies on virtual reality simulations (e.g., Azarnoush et al., 2015), including to assess trainee
performance (Dubin, Smith, Julian, Tanaka, & Mattingly, 2017). Virtual reality is also being used therapeutically, such as to develop emotional skills in children with autism spectrum disorders (Lorenzo, Lledó, Pomares, & Roig, 2016).

**Statistical Literacy**

Many expected changes in the nature of work relate to enhanced access to, and use of, data to make decisions. While some basic statistical processes can be automated with algorithms and machine learning, Gal (2002) argued that citizens must have a basic understanding of the concepts underlying statistical reasoning and terminology to interpret data, evaluate sources, and make decisions, including:

1. Basic statistical and mathematical methods and terminology
2. Foundational statistical concepts, such as probability and variability
3. Inferential reasoning
4. World knowledge, to aid interpretation and evaluation of findings.

Several trends will lead to increased demand for employees in all occupations who can (a) effectively use data, (b) understand how to visualize and manipulate data, and (c) draw conclusions from data. First, data are becoming more accessible. Brynjolfsson and McAfee (2011) pointed out that “Information doesn’t get used up even when it’s consumed…. and once a … body of information is digitized …. it can be copied infinitely and perfectly, and distributed around the world instantly and at no additional cost. This is nothing like the economics of traditional goods and services” (p. 73). The ubiquity of data and the relative ease with which it can be distributed and shared opens the possibility of extensive data analytics. Bonney et al. (2009) discussed the role of such “citizen scientists” in the advancement of both educational and scientific outcomes.

Second, access to new technology has further increased the importance of statistical literacy. Big data, artificial intelligence, and machine learning play an increasing role in daily life. While these technologies will replace some existing human tasks, they are expected to create new skill requirements, such as the ability to train algorithms (e.g., generating models for machine learning used in automated item scoring or image recognition used for automated driving), explain how they work, and keep them operating (Wilson, Daugherty, & Morini-Bianzino, 2017). Statistical reasoning skills will underlie all of these roles.

Finally, statistical literacy can also require specific skills, such as data visualization (Fox & Hendler, 2011). Hampton and colleagues (2017) created a taxonomy of skills for data-intensive research that includes five skill areas (data management and processing, software skills, analysis skills, data visualization, and communication and collaboration for results dissemination). While their focus was on environmental science, this taxonomy applies equally to work in other research-oriented domains as well.

**Computational Thinking**

Computational thinking can be thought of as a special case of analytical thinking, one that draws specifically on the ability of computers to abstract and automate problem solving (Wing, 2008). It is now pervasive in most analytical disciplines (Beheshti et al., 2017).

Hu (2011) defined computational thinking as:
thinking to solve problems, automate systems, or transform data by constructing models and representations, concrete or abstract, to represent or to model the inner-working mechanism of what is being modeled or represented as an information process to be executed with appropriate computing agents. Such thinking is necessarily:

- logical, to capture what is essential to the models or representations;
- algorithmic, to step-wise define or refine operational processes;
- scientific, to gain understanding of models’ capabilities, learn how to use them with maximum efficiency, and explore the effects of the computation in the original problem domain.
- mathematical, to be able to show the correctness of algorithms, specify precisely the functionality of a software system, measure the quality of what we do in a process of computation, and deal effectively with the complexity of the models and representations by exploring more effective and efficient alternatives;
- analytical, to model with purpose, assumptions and viewpoints, evaluate and adjust the models and representations by prototyping, and study their implications and consequences;
- engineering-oriented, to design the models and representations against known constraints and practical concerns, and to plan, execute, manage, and evaluate the process of computation in order to improve our capability and maturity level; and
- creative, to model the unthinkable.

It is important to note that computational thinking is not synonymous with coding or programming skills, but encompasses understanding of computational concepts, practices, and perspectives (Lye & Koh, 2014). Psycharis (2018) argued that computational thinking integrates mathematics, computer science and knowledge in one or more subject areas to solve complex problems. In measuring criterion validity of the Computational Thinking Test, Román-González, Pérez-González, and Jiménez-Fernández (2017) found statistically significant correlations between computational thinking and spatial, reasoning, and problem-solving abilities.

**Interpersonal Skills**

**Communication**

Communication is, at its core, an exchange of information, whether linguistic or non-linguistic, and is widely considered a key competency in both postsecondary education and workplace contexts (Brink & Costigan, 2015). Effective communication, defined as the ability to synthesize and transmit ideas, is among the critical skills needed by employees at all levels of organizations (American Management Association, 2012). Communication in the future world of work will require the ability to work with emerging technologies, along with the more traditional elements of communication such as listening, initiating and engaging in conversation, and persuading others.

**Listening**

Listening has been identified as the most important oral communication skill for successful job performance across a range of workforce samples (Brink & Costigan, 2015). Listening can be categorized into four major types:
1. Active- giving full attention when others are speaking
2. Involved- giving most of one’s attention to the speaker’s words and intents
3. Passive- Receiving information rather than being an equal partner in an exchange
4. Detached- Withdrawn from the speaking-listening exchange such that one is the object of the message rather than the receiver (Pearce, Johnson, & Barker, 2003).

Active listening, in particular, is sought after by potential employers as it helps to create a positive work culture and supports collaboration, which in turn spurs innovation (Nowogrodski, 2015). An active listener fully concentrates on what is being communicated and provides both verbal and nonverbal feedback in response. Examples of verbal feedback include providing positive reinforcement, remembering prior details, asking relevant questions, paraphrasing what the speaker has said, and requesting clarification. Examples of nonverbal feedback during active listening include smiling, making eye contact, maintaining posture, mirroring facial expression, and maintaining focus (SkillsYouNeed, 2018).

**Conversation**

Conversation among team members, whether virtual or face-to-face, is anticipated to be an expanding feature of future jobs (Gribben, Becker, & Dickinson, 2018). Conversation skills are important because they contribute to an organization’s shared understandings, which may be critical for the agile decision-making that is characteristic of the workforce of 2030 (Heidema, 2017). Conversing skill has been rated among the most important oral communication skills, typically rated between listening and presentation in terms of importance for job success (Brink & Costigan, 2015). Employers seek employees with conversation skills because they will contribute positively to the workplace culture by promoting dignity and increasing motivation (Macaulay, 2014).

Conversation goes hand-in-hand with listening, such as knowing when it’s time to listen and when it’s time to talk, and gauging one’s delivery based on mindful listening to the other person’s message (Macaulay, 2014). Conversation skills also include staying organized, and being strategic about the information one both conveys and takes away from the interaction (Coplin, 2003). In the digital age, workers are increasingly engaging in multiple communications, often simultaneously, and therefore run the risk of tuning out important conversations as they seek to filter all of the information received. The ability to engage in authentic conversations on social media platforms therefore becomes a skill of its own (Lombardi, 2014).

**Persuasion**

Persuasion is a uniquely human skill, one that is expected to withstand the wave of increased automation (Luckin, Baines, Cukurova, Holmes, & Mann, 2017) and one that is increasingly in demand (Deloitte, 2016). Also, as the future world of work will be characterized by increasing diversity and geographic dispersion, the ability to persuade people from a variety of backgrounds will be valued in the workplace (Martin, 2010). Persuasion is an important skill for any job role that involves managing customer or client relationships, or managing other employees (Dellaert & Davydov, 2017).

Persuasion skills include (a) making an assessment of the individual or group one is trying to persuade, (b) establishing rapport with them, (c) communicating the benefits of
the proposed course of action, (d) actively listening to any counterarguments, (e) clearly presenting counterpoints to these arguments, (f) recognizing any limitations of the original course of action, (g) modifying the course of action as needed, (h) reaching terms with the person being persuaded, and (i) following up to ensure that they are still on board with the agreed upon course of action (Doyle, 2018).

**Relationship Building**
As the workplace of the future relies more and more on teams, both virtual and in-person, relationship building skills become more valuable. Building relationships has benefits for employees, teams, leaders, and organizations, such as building trust, boosting morale, and improving decision making (Pauleen, 2004). The ability to forge positive relationships in the workplace is key for an individual’s job satisfaction, and is an essential building block in the creation of a collaborative work environment.

Relationship building is characterized by listening to others and encouraging them to share their thoughts and feelings (Lievens & Sackett, 2012). Communication skills therefore play a major role in relationship building, but also key are things such as following through on commitments and being considerate of others’ feelings and perspectives (Tingum, 2018). Relationship building skills are characterized by willingness to share one’s knowledge and expertise; providing quality feedback to others; supporting others’ work while also bringing in others to help with their own work; and engaging in ongoing, friendly interactions inside and outside the workplace (Garfinkle, 2018).

**Cultural Sensitivity**
Cross-cultural competency will be a core skill in most organizations of the future, as employees will need to be able to identify shared values to work effectively with increasingly diverse coworkers (Davies, Fidler, & Gorbis, 2011). In the context of the workplace, cultural sensitivity includes the ability to work effectively alongside someone from a different cultural background who may approach workplace behaviors differently (Sherman, 2018). Culture-based misinterpretations can have implications for the success of collaborative efforts (Blanding, 2012).

Coworkers from different cultural backgrounds may engage in different behaviors and hold different work-related values. For example, employees from individual-oriented cultures may approach work tasks differently than someone from a group-oriented culture (Heggertveit-Aoudia, 2012), which may influence behaviors such as how employees participate in meetings, the amount of time they spend socializing, and whether they provide feedback or otherwise publicly express opinions (Knight, 2015). Increased cultural sensitivity could help mitigate such differences.

At its most basic level, cultural sensitivity requires knowledge and understanding of other cultures (Lutz, 2017). Cultural sensitivity may also involve taking an interest in another culture, recognizing cultural differences, and then changing one’s own behavior to show respect for the other culture (Hammer, Bennett, and Wiseman, 2003). Recognition of one’s own biases is also an element of cultural sensitivity (Loue, Wilson-Delfosse, & Limbach, 2015).

**Understanding Other People’s Perspectives**
Increasing levels of collaboration among diverse teams in the workplace will boost the value of perspective taking as a job skill. Perspective-taking refers to the ability to take on another person’s point of view. It is an active and goal-directed process that involves trying to
understand the thoughts, and feelings of another, as well as the motivations behind them (Parker, Atkins, & Axtell, 2008).

Situational awareness and personal awareness are two key components of perspective-taking. Situational awareness refers to understanding the context in which another person is acting. Personal awareness refers to understanding what the other person brings into that context (Goulston & Ullmen, 2013). Other building blocks of perspective-taking include being aware of others, regulating one’s emotions and empathy, being able to successfully “read” other people, and correctly interpreting what others are trying to communicate (Campbell, 2016).

**Collaborative Problem Solving**
The demand for collaborative problem-solving skills is anticipated to experience high levels of growth in the future. It is defined as the ability to engage effectively with two or more people to solve a problem through shared understanding and effort, and pooled knowledge and skills (Luckin, Baines, Cukurova, Holmes, & Mann, 2017). Collaboration will be key as increasingly complex problems will not be solved by one specific field of expertise, but rather will require working with others from different disciplines (Davies, Fidler, & Gorbis, 2011). Collaborative problem solving was recently added to the skills measured by the Program for International Student Assessment (PISA), a reflection of its significance as a desired skill. As the workplace of the future will be characterized by increasing amounts of teamwork, being able to collaborate to solve problems will be a highly desired skill (Thompson, 2016).

Collaborative problem solving is not only useful for completing job tasks; it is also applicable to maintaining a positive work environment. Managers may use collaborative problem solving to resolve issues among employees by engaging in collaborative discussions to reach a common understanding of the problem at hand and to negotiate a solution (Bernstein & Ablon, 2011).

**Social and Emotional Intelligence**
Emotional intelligence (EI), sometimes referred to as social and emotional intelligence, refers to an individual’s capacity to recognize one’s own and others’ emotions, use this knowledge to inform thinking and behavior, and adapt to meet goals. The concept has had a controversial history since the mid-1990s. Multiple definitions of EI—and various measures associated with each definition—exist today. Some studies have found positive correlations between EI scores and job performance and leadership skills; other studies find no unique contribution of EI to these outcomes beyond correlations accounted for by general intelligence and measures of generally accepted personality traits. We do not delve into the history and nuances of EI here, but instead summarize literature regarding EI’s perceived place among the skills needed in the workplace of the future.

Frey and Osborne (2013) conferred with experts in machine learning to determine the binary likelihood (i.e., yes/no) of automating 70 occupations based on their O*NET characteristics. These occupations were selected from the full suite of 702 O*NET detailed occupations based on confidence in the automation rating. Authors then used statistical modeling approaches to estimate the probability of automating the remainder of the occupations. After extensive analysis, the authors conclude “…as technology races ahead, low-skill workers will reallocate to tasks that are non-susceptible to computerization— i.e., tasks requiring creative and social intelligence. For workers to win the race, however, they will have to acquire creative and social skills” (p. 45).

Using a very different approach, PricewaterhouseCoopers (PwC; 2018) began a collaboration with the Said Business School in Oxford in 2007 to map influential business factors. The study
authors postulated four “worlds of work” to emerge by 2030 in which potential workplace scenarios are described in four quadrants defined by two dimensions: fragmentation vs. integration and collectivism vs. individualism. The authors developed descriptions of each scenario, including a timeline of milestones between 2020 and 2030, major characteristics of the quadrant, implications for workers, what the workforce will look like, and organizational challenges. Following a discussion of all four worlds, they predict the following about jobs: “Automation will not only alter the types of jobs available but their number and perceived value. By replacing workers doing routine, methodical tasks, machines can amplify the comparative advantage of those workers with problem-solving, leadership, EQ (Emotional Intelligence), empathy and creativity skills” (p. 30). PwC commissioned a survey of 10,000 individuals in China, India, Germany, the U.K., and the U.S. and found that 76 percent of respondents agreed or strongly agreed that they had emotional intelligence.

Intrapersonal Skills

**Time Management**

Time management skills encompass a variety of specific abilities: estimation of effort, scheduling, prioritizing, delegation, and monitoring a to-do list, among myriad others. An individual with strong time management skills can not only project the amount of time and effort a given task will require, but also inhabit the mindset to meet deadlines and, perhaps as importantly, recognize when a deadline cannot be met and adapt accordingly.

While employers have historically valued employees with solid time management skills, in the expanding gig economy⁴³ the individual entrepreneur’s personal success depends upon it. When multiple gigs are underway, the ability to schedule and complete each gig as though it was the individual’s only job is critical. In a very dynamic environment where freelancers and hiring agencies are mixed-and-matched in various combinations for specific tasks, the entrepreneur cannot rely upon the understanding of a long-time employer who is familiar with the individual’s work and is sympathetic when projects fall behind. Task matching search engines such as Upwork (www.upwork.com), TaskRabbit (www.taskrabbit.com), or Gigwalk (www.gigwalk.com) collect customer satisfaction data and use this feedback to determine whether to match entrepreneurs to future tasks. Poor ratings due to a lack of effective time management could prevent further assignments.

**Efficiency**

Similar to time management skills, traditionally employers have valued efficient employees. For the 21st annual survey of CEOs worldwide, PwC interviewed 1,293 CEOs in 85 countries, including 104 from the United States, in October and November of 2017 (Ryan, Sapin, Rao, & Ampil, 2018), U.S. CEOs are hiring for broadly relevant digital skills and collaborative, creative, and efficient work styles.

In a gig economy, however, efficiency is particularly important. The individual entrepreneur may face challenges of scalability. At the extreme, the artisan who produces hand-made items can only produce so much; the individual service provider can only manage a limited number of clients or tasks. In order to scale up—which may be necessary in order to obtain a livable wage—the entrepreneur must be efficient.

---

⁴³ A gig economy refers to a labor market characterized by the prevalence of short-term contracts or freelance work as opposed to permanent jobs.
Efficiency can be instantiated in a variety of ways in a gig environment. In some job markets, the individual entrepreneur may add staff and delegate work in order to increase production; however, in the examples just cited—creator of artisanal handmade items or tasker such as a personal shopper or errand runner—adding staff may not be feasible. Alternatively, the worker may use technology to offload mundane or repetitive tasks and free up time for more creative or complex work, requiring human skills. For example, subscribing to a task-matching search engine is an efficient way to seek work, relative to searching for opportunities and applying individually for each. Thirdly, a worker may leverage innovation or creativity to complete tasks more efficiently.

Adaptability

One of the common concerns about the future workplace is that automation will obviate the need for humans to perform large categories of jobs. Certainly, computerization and robots have been demonstrated to be effective replacements for humans in predictable, repetitive environments such as assembly line work. Further, AI has been successfully deployed to rapidly and accurately process large amounts of data to detect patterns and make complex, data-informed decisions. More recently, AI learning systems have been trained to determine optimal ways to conduct certain processes, and monitor their own ongoing effectiveness for further improvement. In this way, automated systems have the capacity for selected adaptability.

Many jobs, however, are less well-suited to automation. Autor and Dorn (2013) note that as computerization has become increasingly affordable, low-skill workers have shifted from routine tasks to the service industry. They contend that these service occupations are somewhat immunized against automation due to their reliance upon a combination of factors including direct physical proximity and flexible interpersonal communication.

In addition to adaptability being key to specific careers, adaptability will also be integral to the projected evolving career path an individual will undertake over the course of a working lifetime. As companies demand upskilling or reskilling, the adaptable employee will be at an advantage. McKinsey (2017) conducted interviews with experts from industry and academia for the April 2017 Digital Future of Work Summit in New York. Experts included professors and executives from NYU, C3 IoT, New America, WorkMarket, LinkedIn, Arena, and McKinsey Global Institute (MGI). MGI partner Susan Lund opined “For young people today, what’s clear is that they’re going to need to continue to learn throughout their lifetime. The idea that you get an education when you’re young and then you stop and you go and work for 40 or 50 years with that educational training and that’s it—that’s over. All of us are going to have to continue to adapt, get new skills, and possibly go back for different types of training and credentials. What’s very clear is that what our kids need to do is learn how to learn and become very flexible and adaptable” (p. 2).

Finally, in a gig economy the successful entrepreneur must be prepared to provide services to multiple employers in a variety of environments. Being able to adapt to technical and administrative requirements will serve the independent contractor well.

Blended Skill Sets

Learning Agility

Learning agility is a skill necessary for the development of skills of the future. It involves many of the skills required for the work of the future, including adaptability, tolerance for ambiguity, communication and listening skills (Eichinger & Lombardo, 2004), but applies them to the development of other skills. DeMeuse, Dai, and Hallenbeck (2010) indicate that learning agility
is related to past experience, self-awareness, and the ability to handle complexity. Mueller-Hanson, White, Dorsey, and Pulakos (2005) relate learning agility to adaptability.

**New Media**

New media refers to the emerging means of communication with large groups of people, and includes the internet, as well as more recent interactive, digital platforms through which more and more people access and consume information (Wynne, 2017). With increasing globalization and dispersion of the workforce, and with a growing number of employers creating and maintaining a new media presence, the ability to effectively navigate in this environment will become an increasingly valued skill (Gribben, Becker, & Dickinson, 2018).

Several core competencies have been identified as essential for participating in new media (Jenkins, Purushatma, Weigel, Clinton, & Robison, 2009). These include:

1. **Play** - the capacity to experiment with one’s surroundings as a means of problem-solving
2. **Performance** - the ability to adopt alternative identities to improvise or discover new things
3. **Simulation** - the ability to interpret and construct dynamic models of real world processes
4. **Appropriation** - the ability to sample and repurpose media content in meaningful ways
5. **Multitasking** - the ability to scan one’s environment and focus in on the key details
6. **Distributed cognition** - the ability to have meaningful interactions with tools that expand mental capacities (e.g., calculator, Wikipedia)
7. **Collective intelligence** - the ability to pool and compare knowledge with others to achieve a common goal
8. **Judgment** - the ability to assess how reliable and credible various information sources are
9. **Transmedia navigation** - the ability to follow stories and information across multiple media
10. **Network** - the ability to search for, synthesize, and distribute information
11. **Negotiation** - the ability to travel across diverse communities, recognizing and respecting differing perspectives, and understanding and following alternative norms
12. **Visualization** - the ability to translate information into visual models and to understand the information that other visual models are conveying

**Projections of Shifts in Future Skills**

**Skills Expected to Increase in Demand**

Literature on trends in workforce skills typically does not include information to determine the rate of the trend. Thus, it is difficult to identify which skills are expected to increase in demand. In the next sections, we present skills that have been deemed critical for the future workforce and show evidence of expected increase in demand. Based on projections of work in the future, there is evidence of expected increase in jobs demanding the following subset of future postsecondary skills.

**Complex Problem Solving**

Futurist Cynthia Wagner (2011) forecasts a limitless supply of future problems to solve. Educators consider critical thinking and complex problem solving a core skill for students (Lara, 2018; P21, 2016) in part because employers increasingly seek workers with complex problem solving abilities.
solving skills. Employers have consistently rated complex problem solving as the most essential competency for career preparedness (National Association of Colleges and Employers, 2017). The ability to address complex problems is particularly important in work environments characterized by rapid change (Middleton, 2002).

**Computational Thinking**
Opportunities to learn how to code, especially for girls (see girlswhocode.com) amidst efforts to achieve gender parity in the computer science industry, have exploded in popularity (Bourque, 2016). Games that teach programming logic are being marketed for children as young as four-years-old (see Kodable.com). Hackathons offer monetary awards and bragging rights, spurring innovative design and problem solving. However, Grover (2018) argues that not everyone will become a computer programmer. Computational thinking – the “ability to translate vast amounts of data into abstract concepts and to understand data-based reasoning” (Davies, Fidler, & Gorbis, 2011, p. 10) – should be taught to every student (Grover, 2018).

**Communication**
Using government job-growth projections, the Pew Research Center identified the fastest growing occupations and skills and preparation requirements for working in those fields (DeSilver, 2016). Nine of the top 10 fastest growing occupations, with projected growth of 5.2% to 13.1% in the ten-year period 2014 to 2024, require job preparation (i.e., formal education, on-the-job training, and prior related experience), interpersonal skills (e.g., communication), or both.

**Cultural Sensitivity and Communication**
According to Davies, Fidler, and Gorbis (2011), ease and sensitivity in working with culturally diverse colleagues will become an important skill for all workers, not just those who work in global corporations. Employers are increasingly recognizing the value of cultural competence and communication skills among new hires (Vozza, 2016), especially when those skills are needed to perform future jobs that involve interaction on a global scale.

**Collaborative Problem Solving**
The demand for collaborative problem-solving skills is anticipated to experience high levels of growth in the future (Thompson, 2016). In recent years, employers have consistently rated collaboration among the top competencies needed for career preparedness, and recognize it as the top attribute for setting apart a potential employee’s resume (National Association of Colleges and Employers, 2017). Industry strategists expect collaboration to be a top business objective in the workplace of the future (Bowles, 2018; P21, 2016). Virtual collaboration (Davies, Fidler, & Gorbis, 2011) will be more in demand as multifaceted problems are too complex to be solved within one discipline or organization (e.g., climate change) and necessitate working on virtual teams. Working on collaborative teams allows for skills gaps to be bridged and increases efficiency (Boyer, 2017).

**Social and Emotional Intelligence**
Davies, Fidler, and Gorbis (2011) declared social and emotional intelligence as one of ten skills critical for success in the future workforce. The Partnership for 21st Century Learning (2016) includes interpersonal skills as a key skill for students to learn before graduation from high school and embarking on postsecondary pathways.

**Adaptability**
Business trends indicate a need for more employees who are comfortable adapting to frequent changes to their work and the workplace (Davies, Fidler, & Gorbis, 2011; McKinsey & Company,
2017; P21, 2016). As organizations are required to respond quickly to changes in an increasingly globalized and technologically advanced world, employers will seek an agile workforce capable of responding to unanticipated change with speed (Breu, Hemingway, Strathern & Bridger, 2002). Workers of the future may be expected to rotate among a variety of roles and tasks (Wadors, 2017). With increasing contract and gig positions expected in the future (Yaraghi & Ravi, 2016), individuals will rely on their adaptability skills as they move from one job to another. Additionally, adaptability skills will facilitate keeping up with changes in technology. Adaptive thinking is a key to innovation and creative problem solving (Davies, Fidler, & Gorbis, 2011).

**New Skills Expected to Emerge**

While information about specific future jobs was scarce, prediction of new skills is non-existent. Using projections of new jobs requiring creativity and social and emotional intelligence, we can extrapolate the need for skills drawing on creativity and social and emotional skills. Increases in artificial intelligence (AI) applications may lead to a need for new skills combining knowledge of AI and other skills to further the research, development, and integration of AI into the workplace.

When futurists discuss workplace skills they tend to focus on combinations of current skills, such as skills in multiple subject-area domains (see transdisciplinarity in Davies, Fidler, & Gorbis, 2011) or skill in cognitive and non-cognitive areas (e.g., analytical and interpersonal skills). By 2030, some new jobs will likely demand completely new skills that we cannot yet imagine and therefore currently cannot describe or label.

**Summary of Themes of Skills for the Future**

When it comes to skills for the future, the types of jobs offered and changes to workplace processes will be the major drivers of skills needed to enter the workforce. By extension, skills for jobs needing postsecondary education or training require skills for success in school as well as job-specific and cross-career skills. Although change is expected in the world of work, the amount and direction of change is unknown. However, this review of relevant literature points to some overarching themes that provide a solid base for making predictions about career skills that students who entered kindergarten in 2017 will need in 2030 when they graduate from high school. Careers will likely look very different from those their parents were prepared for, particularly in terms of the number of jobs and variety of skills needed across and within those jobs over their lifetime.

Jobs of the future like those of today will require a mix of cognitive, interpersonal, and intrapersonal skills. The specific set of skills and which ones are most important will vary depending on the pathway a student follows. Increasingly, blended skill sets (e.g., learning agility and new media) will be needed. Fields that had previously been quite separate may be blended in new ways, requiring combinations of skills not seen before. Existing jobs may be blended with new technologies to create positions we have never seen (think: space junk recyclers) and requiring new skills or blended skill sets.

Industry believes “data is the new oil.” To meet the expected rise in demand for employees to work with data, employers will seek individuals with skills related to manipulating, analyzing, interpreting, and illustrating data.

High school graduates of 2030 will set out on career pathways characterized by change. Whether they work independently through the gig economy, or move among multiple employers or across multiple departments or projects, workers of the future will likely need to be adaptable and agile as they will be asked to respond quickly to unanticipated changes. They will need to
draw upon cross-cultural and virtual collaboration skills as they find themselves part of an increasingly diverse and dispersed workforce. Workers will become lifelong learners as jobs continually evolve to meet changing demands and to incorporate the latest innovations.

With some sense of what the future holds, a key next step to ensuring that students graduate high school in 2030 prepared for success in their postsecondary pathways is to measure the skills needed to perform the jobs of the future. Assessing future postsecondary skills provides a metric for understanding and monitoring how prepared the next generation is for life after high school.
References


Land, M. H. (2013). Full STEAM ahead: The benefits of integrating the arts into STEM. *Procedia Computer Science*, 20, 547-552


Thompson, C. (2016, January 21). The top 10 skills that will be in demand by all employers by 2020. Business Insider.


A Review of State Indicators of College and Career Preparedness

Introduction

The purpose of this memorandum is to review states’ practices for measuring college and career readiness. The memorandum begins by exploring how states have defined the constructs of “college and career readiness.” Next, common and well-established measures such as college entrance exams and statewide assessments in English Language Arts (ELA) and mathematics are described. Then, some less-established academic indicators are discussed. Finally, the memorandum discusses other career or work readiness indicators, including non-academic and other innovative indicators, that states may be including in their suite of exams to tap into career or workplace skills.

The information included in this memorandum was obtained from (a) states’ federal peer review submissions, (b) publicly available information on websites (e.g., state websites, research organization websites, test vendor websites, governmental organization websites), and (c) a review of current articles and scientific literature (e.g., EdWeek articles). Each listed indicator/assessment includes its own research base and documentation; it is beyond the scope of this memorandum to include all of that information. However, enough detail is provided to allow interested parties to discover more information about assessments of interest.

As this memorandum will illustrate, assessments specifically designed to address “career readiness,” including non-cognitive, social-emotional learning, and career exploration, are not widely used by states. Some of the indicators described in herein may be used by a single state or may not be administered to all students or even to all schools/districts within a state. However, many states are in the midst of procuring new statewide assessment contracts and the landscape of assessment is ever-changing. Thus, there are likely additional indicators of preparedness for college and career that are not included in this memorandum. As such, this memorandum represents a preliminary effort to catalog some of the innovative ways that states are beginning to approach college and career preparedness.

Framing the Need to Address Readiness for College and Career

Preparing students for post-graduation opportunities has long been a priority for states, districts, schools, teachers, and parents. The current expansion of the global economy and ongoing labor market shifts has refocused attention on the readiness requirements for student success in an ever-changing post-secondary landscape. The current workforce demands employees who can fill jobs that require a combination of technical and academic skills (Carnevale, Smith, & Strohl, 2010). However, employers struggle to find such qualified workers (Organization for Economic Co-operation and Development, 2013). Educational institutions struggle with defining and measuring the knowledge, skills, abilities, and dispositions necessary for success in the current post-secondary world. These efforts are key to guiding educators’ efforts to identify which students are on track to succeed as they graduate from high school.

Porter (2018) provided an overview of the current landscape of occupations in the United States and their requirements during a talk at the National Conference on Student Assessment (NCSA). He described ways that states have attempted to upgrade their tests to prioritize the most relevant knowledge and skills for students. The summary figure is reproduced below.
According to Mishkind (2014), 37 states defined college and career readiness as of 2014. For 33 of those, a single definition was used to describe both college and career readiness. The rationales for using a single definition tend to include views on the interconnectedness and similarities between visions of preparedness for college and preparedness for career. For example, the Oregon Investment Education Board\textsuperscript{44} indicates that, “the overarching skills and strategies required for students of all ages entering colleges and careers are consistent.” New Hampshire, in their ESEA flexibility request\textsuperscript{45} stated, “Evidence and experience indicate the knowledge and skills needed to succeed in college and career are greatly similar, and that all graduates will need some form of postsecondary education or training to succeed during their careers.”

The U.S. Department of Education (ED) currently indicates that “Nearly every state now has adopted college- and career-ready standards.” This quote refers to college and career readiness specifically in the subjects of mathematics and ELA. ED also describes the requirements for approval of state assessment and accountability systems under ESSA related to standards, indicating that “All states approved for ESSA flexibility have engaged in one of the following endeavors to raise expectations for students’ academic performance:

- Upgraded their existing standards to make them more rigorous by working with their four-year public universities to certify that mastery of standards ensures that students

\textsuperscript{44} See http://stand.org/oregon/OEIB.
will not need to take remedial coursework upon admission to a postsecondary institution in the system; or

- Adopted and implemented common standards developed by a consortium of states that build toward college and career readiness.\(^{46}\)

According to ED, as of May 31, 2018, 33 states (plus the District of Columbia) included a measure of college and/or career readiness in their ESSA plans.\(^{47}\)

In states where there are separate definitions, there is a tendency for considerable overlap in the definitions. For example, the Nebraska Department of Education defines a career ready person as follows, "A career ready person capitalizes on personal strengths, talents, education and experiences to bring value to the workplace and the community through his/her performance, skill, diligence, ethics, and responsible behavior."\(^{48}\) The Nebraska Career Readiness Standards indicate that when students are career ready, they are prepared for the next steps in their lives—whether that means getting their first job or beginning their college 'career' (which eventually leads to the workplace as well). Being career ready also means being ready for life." Nebraska is one of an increasing number of states to also include career and technical education as part of their statewide accountability system. They provide programs outside of the regular school day for students to connect with resources to develop interests and skills for future success.

Mishkind (2014) identifies specific components of the states’ definitions of college and career ready and places them into the following six categories:

- Academic knowledge (e.g., English, mathematics, core subjects, preparation to take credit-bearing courses in related subjects)
- Critical thinking and/or problem solving (e.g., reasoning, problem solving, analysis, evaluation, presentation of ideas and information)
- Social and emotional learning, collaboration, and/or communication
- Grit/resilience/perseverance
- Citizenship and/or community involvement
- Other additional activities

Distribution of these components among the 37 states identified by Mishkind (2014) is presented in Figure 2. Similar reports reflecting the state definitions under ESSA have not yet been published.

This memorandum treats “readiness” and “preparedness” synonymously, although states may make distinctions between them. Treatment of the terms is not consistent from state to state, with some referencing “readiness” only as a link to college placement tests’ established readiness benchmarks, while others treat readiness more generally. This memorandum is concerned with documenting the ways that states address readiness, so both terms were used in its preparation and no distinctions were made between them.

\(^{46}\) See https://www.ed.gov/k-12reforms/standards.
\(^{47}\) See https://www.ecs.org/50-state-comparison-states-school-accountability-systems/.
\(^{48}\) Definition adopted in 2010, see https://www.education.ne.gov/nce/careerreadinessstandards/.

---

Appendix K. Literature Review: State Indicators of College and Career Preparedness
Typical Measures of Readiness

While states are exploring expanded definitions of readiness, their federal peer review submissions indicate that, for the most part, academic or college readiness is the aspect that is measured. Academic readiness is typically measured by traditional assessments, and work or career readiness is considered either synonymous or is described as a set of traits or characteristics that extends beyond academic readiness. States assume academic readiness is a necessary component of preparation for both college and career. Most states do not include measures of non-academic readiness in their accountability systems.49

1. State Accountability Tests

State are required under ESSA to test students in grades 3-8 and high school in ELA and mathematics. Some statewide assessments are administered on computers while others are administered via paper and pencil. These assessments are typically a mix of multiple-choice and short constructed response (students provide a brief written response) items. Some state tests also include multiple-select items (which are selected response items with potentially more than one correct answer), essay or extended constructed response items (longer writing pieces), or technology enhanced items (TEIs). TEIs must be computer administered and may include matching, drag-and-drop, ordering, or other tasks that are more readily performed in the computer administered environment.

49 See https://www2.ed.gov/admins/lead/account/stateplan17/statesubmission.html, for a full description of states’ peer review submissions.
Most statewide summative assessments are based on the Common Core State Standards for (CCSS) ELA and mathematics. The figure below, from the Association for Supervision and Curriculum Development (ASCD),\textsuperscript{50} shows the state level adoption of the CCSS as of 2018. States highlighted in green are adopters, Minnesota, (highlighted in blue) adopted the ELA but not mathematics standards. States highlighted in gray did not adopt the CCSS. This graphic shows there is more agreement among states than not regarding what academic content students should learn in school, at least for ELA and mathematics.

![Figure 3. CCSS adoption, by state.](image)

In addition to largely agreeing on tested content, states have been steadily moving toward asking students to perform more complex and challenging tasks on summative tests. The testing consortia, Smarter Balanced Assessment Consortium and the Partnership for Assessment of Readiness for College and Career (PARCC), require a great deal of strategic thinking (critical thinking, reasoning, and developing a plan) and even extended thinking (thinking to investigate a problem and synthesize information) in a substantial portion of their assessment items. Figure 4 compiles information on the level of complexity among several summative state tests (Porter, 2018). This research was compiled from studies conducted by HumRRO (Schultz, Wiley, Michaels, & Dvorak, 2016), the Thomas B. Fordham Institute (Doorey & Polikoff, 2016) and Rand (Faxon-Mills, Hamilton, Rudnick, & Stecher, 2014).

States have adopted challenging content standards and state assessments have become increasingly more complex. However, comparability of students’ scores across states remains challenging. Unless states use a common assessment with common proficiency definitions (such as those states in consortia), students’ proficiency may depend on the state where they were tested. To address comparability, states participate in reading and mathematics testing on the National Assessment of Educational progress (NAEP).\(^{51}\) NAEP is a complex suite of assessments based on rigorous academic frameworks documents. NAEP tests do not report individual student scores, but report at the state level based on samples of students. The National Center for Education Statistics (NCES) produces maps between state and NAEP proficiency standards, allowing states to compare the rigor of their standards using a common assessment.\(^{52}\) States often reference the NAEP Mathematics and Reading Frameworks and the NAEP proficiency standards when they design their own assessments.

### 2. High School GPA

Student academic performance is also tracked by grade point average (GPA), a simple transformation of the letter grades they receive for their courses. GPA is not typically tracked at the state level as part of accountability, but grades are commonly used as an indicator of college readiness and GPA is a required component on most college applications. Some states, notably Kentucky, use grades to determine supplemental funding provided to help students pay for college. The Kentucky Educational Excellence Scholarship (KEES) program rewards students who graduate with at least a 2.5 GPA, and the amount of money provided to students goes up if their GPA is higher.\(^{53}\)

The measure of GPA includes more than academic knowledge (Farrington, Roderick, Allensworth, Nagaoka, Keyes, Johnson, & Beechum, 2012). The requirements for achieving strong grades in classes also incorporates perseverance, compliance, and time management. Depending on the nature of their classes, students may also need to demonstrate their ability to

---

51 See [https://nces.ed.gov/nationsreportcard/](https://nces.ed.gov/nationsreportcard/).
52 See [https://nces.ed.gov/nationsreportcard/studies/statemapping/#](https://nces.ed.gov/nationsreportcard/studies/statemapping/#).
53 See [https://www.kheaa.com/website/kheaa/kees?main=1](https://www.kheaa.com/website/kheaa/kees?main=1). Students can increase their KEES scholarships with high test scores as well (e.g. ACT, Advanced Placement, International Baccalaureate, Cambridge Advanced International).
collaborate with others, communicate their ideas, and even to construct products. Grades are more far-reaching than summative tests. However, grades are not standardized from teacher to teacher, school to school, or district to district, and they are not typically used by states in their accountability computations.

3. Course Requirements/Completion

States require that all students complete a certain number of courses in specific subjects prior to graduation. These required courses include ELA, mathematics, science, and social studies; however, they may also include arts, health, physical education, foreign language, or others. The Education Commission of the States (ECS) maintains a website that details all states course requirements for graduation.54 In addition to state requirements, districts may add their own requirements for earning a high school diploma.

4. End-of-Course Assessments

States may use either summative tests for high school students administered in particular grades, or they may tie testing to the completion of courses, or both. When states choose to use end-of-course (EOC) exams, they are most often tied to English 10 or 11, and to either Algebra 1 or Algebra 2 courses. For example, PARCC created assessments for multiple English and mathematics courses for states, including subject-specific mathematics courses like Algebra and Geometry, as well as Integrated Mathematics. These types of tests limit the amount of content to be assessed to that addressed by a single course (as opposed to the full range of high school standards), but they assume all courses with a common name cover essentially the same content. States may require or encourage schools to use the assessment scores as part of student’s grades, or as a course completion requirement. EOC assessments may be used for state accountability computations.

Even states who use EOC exams vary greatly on the number of EOC exams offered, the courses for which they are offered, and whether EOC exams are used for accountability. For example, a state may use an Algebra 2 EOC exam for high school accountability, but offer an Algebra 1 EOC exam that is not used. For states with science accountability in place for high school, this is further complicated by course taking patterns (science courses do not follow a particular sequence and students may not take certain science courses). Some states provide EOC exams for Biology for accountability, and require that all students take a biology course, while others require a more general science exam administered at a particular grade level. Others require the Biology EOC exam, but supplement it with other science content to address physical and earth/space sciences.

5. Graduation Exit Exams

The following thirteen states used high school graduation (or exit) exams as of 2017; Florida, Indiana, Louisiana, Maryland, Massachusetts, Mississippi, New Jersey, New Mexico, New York, Ohio, Texas, Virginia, and Washington55. Exit exams require that students demonstrate some minimum level of academic competency, typically in ELA and mathematics, as a condition of graduating. Students who do not pass the exams do not receive a high school diploma. While not specifically designed to indicate college or career readiness, high school exit exams help certify to employers that graduates can be expected to have a minimum level of competency in

54 See www.ecs.org
55 See https://www.fairtest.org/graduation-test-update-states-recently-eliminated
the tested subjects. These types of exams are controversial for several reasons, including the argument that they may reduce graduation rates and increase dropout rates, and they may test content that is commonly found in middle school standards rather than in high school. Exit exams tend to focus on literacy and numeracy, rather than the more complex high school standards. The literature is mixed on these topics and too extensive to adequately review here. Arkansas, Arizona, Georgia, Idaho, Minnesota, Oklahoma, Rhode Island, and South Carolina had graduation exams, but they recently terminated the requirement for high school graduation. California, Nevada, and Pennsylvania currently have a moratorium on graduation tests.

6. College Entrance Exams

College entrance exams are another commonly used type of assessment of readiness for specific courses in colleges or universities. Common examples of entrance exams include College Board’s Accuplacer assessments, ACT’s Engage assessments (ACT previously promoted a placement test called Compass), and a multitude of college- and university-created placement tests. While not typically used statewide, these assessments provide indications for higher education of the likelihood that students will successfully complete specific courses. Results from these assessments may be used to place students into non-credit remedial courses at a college or university.

In the early years of its reporting as part of state accountability systems, college entrance exams were frequently used as an indicator of college and career readiness. For example, the Southern Regional Education Board (SREB) presented changes in average SAT composite scores as an indication of trends in college and career readiness (SREB, 2012). As the concept of readiness for college and career has become better articulated, it is clear college entrance exams may not be the most appropriate tool for measuring career readiness.

The ACT and SAT programs are very clear about their intended purposes. The ACT is designed to measure “academic readiness for college.” Similarly, the SAT measures what students “need to succeed in college.”

Approximately half of states’ accountability systems incorporate college entrance exams as an indicator of college and career readiness. Typically, student performance (e.g., meeting established benchmarks) is measured, but in a small number of states only rates of participation in these exams is reported. In no states are college entrance exams the sole indicator of college and career preparedness. Typically, these scores are reported along with several other indicators, such as performance in advanced coursework (e.g., AP or IB classes), as well as performance on WorkKeys®, earning credentials or certifications, and/or performance in career and technical education classes (American Institutes for Research, 2018).

College entrance exams have the benefit of large, well-established bodies of validity evidence, and as such may be viewed as a cost-effective resource for measuring postsecondary preparedness. However, one potential concern is the extent to which the content knowledge and skills measured by these exams aligns with the content standards adopted in accordance with state law. Most states have adopted standards that are closely related to the CCSS. ACT

56 See https://accuplacer.collegeboard.org/
Appendix K. Literature Review: State Indicators of College and Career Preparedness

(2010) conducted an evaluation of the alignment between its College Readiness Standards and the CCSS and found substantial levels of alignment in terms of reading and mathematics, with weaker alignment in writing. Similarly, the College Board found strong alignment between the SAT and common core anchor standards in reading, writing, language, and mathematics (Vasavada, Carman, Hart, & Luisier, 2010). However, a recent study by Assessment Solutions Group found it would not be appropriate to allow school districts in Florida to opt to use the ACT or SAT in lieu of its state assessment (Roeber, Olson, & Topol, 2018). Currently, 25 states require all high school students to take the ACT or SAT, and 12 states use the ACT or SAT for accountability.61

College entrance exams such as ACT and SAT identify cut scores, or benchmarks, to assist in defining and evaluating student progress toward, and achievement of, adequate levels of college preparedness. States may be required by law to identify benchmarks specific to their own state. For example, the Wisconsin Department of Public Instruction convened panels of educators and other stakeholders to evaluate ACT benchmarks and determine the appropriateness of identifying state-specific benchmarks (Wisconsin Department of Public Instruction, 2015).

**Academic Readiness Measures Beyond ELA and Mathematics Achievement**

Academic preparation for college and career involves more than preparation in ELA and mathematics. High school students take at least 20 courses in grades 9–12, and only eight of those courses are typically required to be English or mathematics courses.62 Students also take courses in science and social studies. They may take classes in the arts, technical courses, foreign language courses, and others. The NCLB and ESSA helped create a focus on English and mathematics, but now states have largely adopted the Next Generation Science Standards (or similar state-specific, three-dimensional science standards) and many have also adopted social studies standards. Many states have begun to test these subjects as part of their accountability systems.

### 1. Science Assessments

Nearly two-thirds of U.S. students live in states that have education standards influenced by the *Framework for K–12 Science Education*63 and/or the *Next Generation Science Standards (NGSS)*64. Figure 5 shows these states in blue and yellow (blue states have adopted NGSS, yellow adopted their own standards based on National Research Council [NRC] frameworks). Both documents promote a more complex phenomenon-based view of science and stress students’ capacity to use their science knowledge in unfamiliar contexts. They address science and engineering principles, cross-cutting concepts, and content knowledge and expect those ideas to intersect for students through rich science instruction.

---

61 For a complete list of states requiring ACT or SAT see https://www.edweek.org/ew/section/multimedia/what-tests-does-each-state-require.html.
62 See www.ecs.org
63 See https://www.nap.edu/read/13165/chapter/1.
64 See https://www.nextgenscience.org/.
Nineteen states and the District of Columbia (representing over 36% of U.S. students) have adopted the Next Generation Science Standards (NGSS). The 19 states are Arkansas, California, Connecticut, Delaware, Hawaii, Illinois, Iowa, Kansas, Kentucky, Maryland, Michigan, Nevada, New Hampshire, New Jersey, New Mexico, Oregon, Rhode Island, Vermont and Washington.

Twenty states (representing 29% of U.S. students) have developed their own standards based on recommendations in the NRC Framework for K-12 Science Education. The 20 states are Alabama, Colorado, Georgia, Idaho, Indiana, Louisiana, Massachusetts, Mississippi, Missouri, Montana, Nebraska, New York, Oklahoma, South Carolina, South Dakota, Tennessee, Utah, West Virginia, Wisconsin, and Wyoming.

Figure 5. State science standards summary.

Most states test science once in elementary school (usually at grade 4 or 5), once in middle school (usually in grade 7 or 8), and once in high school (usually in grade 11 or as an end-of-course biology assessment). Science assessments are only now catching up to the adopted science standards. Some states (e.g., Kentucky) have implemented highly complex phenomenon-based assessments for science (similar to scenario-based tasks used in NAEP). These assessments are similar to passage-based ELA assessments, where students are given a passage to read and then answer questions requiring them to interpret or interact with the passage. In science, students would be given a thorough description of some scientific phenomenon, which may include experimental results, data tables, graphs or charts, or other information referencing the problem the students are expected to help solve. Students demonstrate their understanding of science content by interacting with the phenomenon. Phenomenon-based assessments can take longer to develop, may be more expensive than more traditional tests, and may take longer for students to complete. States are struggling with
the structure of science assessments, how to ensure alignment to complex three-dimensional standards, and how best to report science scores for students, as well as for schools.

2. Social Studies Assessments

Historically, states have not adopted common social studies standards or tests of social studies. While there is a growing trend for states to adopt social studies standards (see above), assessments of social studies remain less common than ELA, mathematics, or science. Social studies assessments, when administered, are typically given at the same grade levels as science. The content of state-wide social studies tests is based on standards that vary considerably from state to state. Some states use EOC assessments of US History in high school, while others test social studies more broadly at specific grade levels.

The National Council for Social Studies (NCSS) has published social studies standards, but they have not been widely adopted by states or included as the basis for assessments in state accountability systems. The NCSS has identified five disciplines of social studies: history, geography, civics and government, economics, and psychology. Those disciplines address ten themes of social studies, including:

- Culture
- Time, continuity, and change
- People, places, and environments
- Individual development and identity
- Individuals, groups, and institutions
- Power, authority, and governance
- Production, distribution, and consumption
- Science, technology, and society
- Global connections
- Civic ideals and practices

Career or Work Readiness

There is growing acceptance of the notion that readiness for college and readiness for career are distinct but overlapping constructs. Patelis (2018) notes there are sets of skills at the intersection of career and college preparedness. Similarly, ACT has developed an integrated framework that defines the core academic skills for readiness from kindergarten through one’s career. The common thread here is that there are key academic components that are common to postsecondary readiness in general, whether a student opts to continue learning in a postsecondary educational setting or in a more applied, workplace-based setting.

Career preparedness includes three major skills needed to perform work: (a) core academic skills; (b) employability skills; and (c) technical, job-specific skills (ACTE, 2010). Academic skills include foundational academic knowledge in mathematics, ELA, and science or technology, including the application of these skills, which were described in the prior section of this memo. Employability skills are skills that are critical to workplace success and include critical thinking, adaptability, problem solving, oral and written communication, collaboration and teamwork, creativity, responsibility, professionalism, ethics, and technology use. Technical skills represent job-specific knowledge and skills needed to enter a career.

Next, some indicators of career readiness that are being used or being considered by states and districts are presented. Because many states are currently exploring career readiness indicators, due at least in part to expanded flexibility in ESSA for using such measures, there is not clear documentation currently available on which districts within which states are using these indicators.

### 1. Assessments of Career or Work Readiness

The Center on Education Policy (2013) found via survey that 36 of 46 participating states assessed students on academic skills related to career readiness. Smaller numbers reported assessing students on technical skills (33 states) and employability skills (28 states). They found the tests most commonly used to assess employability skills and/or applied academic skills were the Armed Services Vocational Aptitude Battery (ASVAB) and WorkKeys® assessments. The ASVAB tests are designed to measure verbal, mathematics, science and technical, and spatial domains. WorkKeys® “measure foundational skills required for success in the workplace” via several assessments including applied mathematics, graphic literacy, workplace documents, applied technology, and business writing, among others.66

A smaller number of states reported administering the National Occupational Competency Testing Institute (NOCTI) assessments. NOCTI offers a Workplace Readiness assessment to assess employability skills at the high school level, as well as tests of employability skills that are targeted toward middle school students.67 Currently, 11 states use NOCTI in some capacity, but their use of the assessments is varied. Some require full census testing in high school, while others make the NOCTI assessments available for school or district use.

In fall 2017, the Center for Educational Testing and Evaluation (CETE) conducted its first operational administration of the Career Pathways Assessment System (cPass® assessments), which are designed to measure “both the knowledge and skills needed for specific career pathways.”68 In addition to a general Career and Technical Education (CTE) assessment, cPass® offers assessments of Comprehensive Agriculture; Power, Structural, and Technical Systems; Animal Systems; Plant Systems; Horticulture; Manufacturing Production; Comprehensive Business; Finance; Accounting; and Drafting. Currently, Kansas and Colorado are members of the collaborative that oversees administration of the cPass®.

Still other states have developed their own career readiness assessments. For example, Kentucky developed the Kentucky Occupational Skills Standards Assessment (KOSSA), which

---


68 See [https://careerpathways.us/](https://careerpathways.us/).
assesses a combination of academic, employability, and occupational skills specific to each of more than 30 career pathways.\textsuperscript{69}

2. \textit{Industry Certifications}

Achieve, Inc. found substantial variability of states’ approaches to incorporating industry certification into educational policy and processes (Muller & Beatty, 2009). Some states guide students along a CTE pathway, ensuring they complete the course requirements that will best prepare them for achieving certification within their preferred industry. Other states allow districts to develop approaches that meet their specific needs. Among the more innovative programs are the tiered diploma system developed in Florida that offers a credential-based graduation option and the “Core 40 with technical honors” diploma offered in Indiana.

3. \textit{Descriptors of Traits or Characteristics}

While not written into statute or as part of the accountability computations, some states have published traits or characteristics of college and career ready students that go beyond academic expectations. These traits or characteristics may include citizenship, financial responsibility, goal setting, problem solving, and reasoning, among others. For example, in 2013, Michigan produced a list of Characteristics of Career and College Ready Students. Michigan’s source documents include the Michigan Literary Standards, National Research Council (NRC) Science and Engineering Practices, Michigan Social Studies Grade Level Content Expectations (GLCE) and High School Content Expectations (HSCE), Michigan Mathematics Practices, Michigan Visual, Performing and Applied Arts (VAPAA) Guidelines, and Career and Technical Education (CTE) Career Ready Practices. Under each source document, Michigan then lists the specific characteristics gathered from each source, grouped into four main headings (a) Technology and Tools, (b) Argument and Reasoning, (c) Communication and Collaboration, and (d) Problem Solving.\textsuperscript{70} Other states have addressed attributes of readiness in similar ways (e.g., Kentucky, Nebraska). According to ED, six states plan to use an art access/participation, or well-rounded education measure, in their accountability systems.\textsuperscript{71} States’ plans tend to focus on the school level, measuring access or participation rates, rather than on individual student’s preparedness.

4. \textit{Social and Emotional Skills Assessment}

Social and emotional skills assessment is a relatively new field for estimating students’ academic preparedness. The ACT Tessera\textsuperscript{72} is focused on the holistic growth of the “whole child.” It measures students’ social and emotional learning skills and provides data to help schools support goal setting and promote continuous improvement. ACT Tessera includes indicators of grit, teamwork, resilience, curiosity, leadership, and school climate. It uses a combination of self-report and forced choice and situational judgment test items to ensure accurate reporting and minimize student “faking.” Reports are provided at the student and school levels, and both students and schools receive recommended resources and strategies to address areas requiring increased focus.

\textsuperscript{69} See \url{https://education.ky.gov/CTE/kossa/Pages/default.aspx}
\textsuperscript{70} Downloaded July 3, 2018 from \url{https://www.michigan.gov/documents/mde/Characteristics_of_Career_College_Ready_Students_v4.11.13_417952_7.pdf}
\textsuperscript{71} See \url{https://www.ecs.org/50-state-comparison-states-school-accountability-systems/}
\textsuperscript{72} See \url{https://pages2.act.org/Tessera-Brochure-Learn-More.html#_ga=2.261599630.22546774.1532353457-1600368859.1524168013}
5. Service Learning/Applied Experience

Service learning, for which students engage in projects or activities to improve their communities or address social problems where they can apply their learning, is increasingly reported as a requirement for high school graduation. To date, only Maryland and Washington D.C. have statewide service learning requirements; however, 19 other states allow districts to require service learning as a graduation requirement. A full description of service learning requirements for graduation by state can be found at the Education Commission for the States website (http://ecs.force.com/mbdata/mbquest3RTE?Rep=SL1301).

6. Interest or Suitability Inventories (Career Exploration)

Although they do not require interest or suitability inventories for career exploration, many states indicate these tools are available and they may provide links to them for interested students. Their use may also be promoted at the district or school levels. Some of these tools provide information at the classroom or school level, but the information provided varies greatly from one tool to the next. In addition, many colleges and universities encourage students to complete interest inventories or surveys prior to applying for admission or choosing a major. They may promote this activity through interaction with high school guidance officers or others with access to high school students nearing graduation.

When available, survey results are used to help students explore the fit between their interests or traits and potential future jobs or college major fields of study. These types of surveys have been used for a long time and there are a wide variety currently available to students. Some of the oldest are based on personality traits, with the personality traits matched to specific careers or job categories. Examples include the Myers-Briggs Type Indicator (MBTI) (Myers & McCaulley, 1985), Keirsey Temperament Sorter (1998), and Strong Interest Inventory (Donnay, 1997). These indicators of traits and their relation to jobs have been used for more than 50 years, with the Strong Interest Inventory used since 1927. The Five Factor Model (FFM) of personality traits is more commonly used today in business and industry, and it has been linked to academic performance indicated by grade-point-average (GPA) in high school and college (Poropat, 2009). These types of inventories assume the traits or interests are not (or are very minimally) mutable. They focus on using information about the students to match them with careers or job families that are best suited to their specific characteristics.

Below are some specific interest inventories/platforms that are currently in use.

6a. ASVAB Career Exploration Program (CEP)

A comprehensive career planning program, the ASVAB Career Exploration Program (CEP) is the only career planning resource available to high schools nationwide that offers high school students the opportunity to explore all paths to career satisfaction—education, work, training, military, and/or certification—because the ASVAB CEP assesses academic ability and career interests.73 The ASVAB CEP first helps students identify skill strengths based on their ASVAB multiple-aptitude skills test results and then links those skill strengths with an interest inventory that highlights work-related interest areas where students are most likely to succeed. ASVAB’s Find Your Interests Inventory is based on Holland’s occupational codes: Realistic, Investigative, Artistic, Social, Enterprising, and Conventional (RIASEC).74 The Holland Codes refer to a theory of career choice based on personality type. Finally, students access a catalog of careers with

---

73 See https://www.asvabprogram.com/
74 See https://www.hollandcodes.com/
relevant occupational data and a full suite of future-oriented planning tools that help them develop an action plan to share with parents and educators.

Equipped with an understanding of their skills and interests, students can explore careers that coincide with their skill-interest profile and their aptitudes. Participation is voluntary. Students in grades 10–12 and college are encouraged to use the ASVAB CEP to identify their skill strengths and explore potential careers that require those skills.

Students gain access to career exploration tools after they receive their scores. Then, they can independently explore any careers that interest them, including those found in the military. The Department of Defense (DoD) sponsors ASVAB CEP with a two-part mission: to provide a career exploration service to U.S. youth and to provide qualified leads to military recruiters. Participants have no obligation to military service, but 11th–12th grade students who choose can use their scores to explore enlistment. ASVAB CEP is provided to students at no cost. Currently, 12 states use military testing or enlistment as an indicator for accountability in their ESSA plans. Kentucky, for example, provides students multiple options for demonstrating postsecondary readiness, including a military readiness route that includes ASVAB and military enlistment.

6b. Career Cruising
Career Cruising’s (https://public.careercruising.com/en/) online interface engages students in exploring their interests and aspirations through interactive games starting in K–2, followed by career activities and assessments in later grades. Career Cruising connects the real world to the classroom with current career and labor market information, salaries, and educational pathways. With this resource, students can explore skills like financial literacy and goal-setting through video interviews, role-playing activities, and more. Using Career Cruising for K–12, students discover how their strengths and interests align with careers that inspire them, and build the academic plan to get them there. Career Cruising helps students understand how the schoolwork they do today impacts the life they live beyond graduation.

Career Cruising for K–12 starts by engaging elementary-aged students through a light-hearted yet informative career education program. There are two interactive game experiences, one for grades K–2 and one for grades 3–5. Students are introduced to many careers through animated characters and solve career-related mysteries as they progress through a fictional town, learning as they go. The experience results in students being more informed about different careers and gaining a basic understanding of academic and career planning, which sets the stage for them to transition to the full Career Cruising program as they move to middle and high school.

Students begin with comprehensive assessments to reveal their own strengths and interests. Then, from the database of hundreds of careers, a select list of careers is presented to the student matching his/her strengths and interests based on his/her answers. Students explore those career possibilities through multiple pathways, including interviews, job descriptions, salary expectations by region, and more. A data library gives students access to college and technical school details, including applications, tuition, enrollment, and more. Students can see which courses will best prepare them for their desired path. Customized for specific graduation requirements, course planning means that students pick their classes in alignment with their career and postsecondary goals. Students can record their goals, save preferred jobs, and track college applications in a digital portfolio. Students make their choices by test-driving career options to see how choices they make might impact their future. Financial literacy lessons help them understand the importance of financial planning and role-playing activities help pull together everything they learn.
6c. Additional Career Interest Inventories and Planning Tools

There are several other career interest inventories and career planning tools on the market that can and are being used by schools. A few examples include Kuder,75 Naviance,76 World of Work Inventory,77 and ACT Engage.78 Increasingly, career interest exploration begins as early as kindergarten and is typically provided in an online environment (see World of Work and Career Cruising).

Conclusions

Most states have adopted a definition of college and career readiness and require schools to test students using measures of readiness as part of their statewide testing programs. Definitions of readiness are typically encompassing and do not differentiate between college and career, although there is commonly a focus on academic preparation for college. Typically, states measure ELA and mathematics achievement and use those scores as the main indicators of readiness.

When states do address career paths specifically, they typically rely on WorkKeys® or narrowly defined industry credentials. Other aspects of career or work readiness, when they are addressed at all, are measured using participation rates or information about access at the school level. There is very little information generated at the student level specific to career readiness in most states.

States are increasingly demonstrating that they value social and emotional learning, but there are no widely used large-scale measures of non-cognitive skills. According to a panel of state education leaders who met at the Council of Chief State School Officers’ National Conference on Student Assessment in 2018, these skills are viewed as essential for both college and career. Skills in communicating, collaborating, creative problem solving (individually or on teams), among others, were listed as essential tools for success in a multitude of careers. The panel also indicated that these skills are very difficult to define and measure.

States are making progress towards addressing college and career readiness, but to date, their efforts have been narrow and focused on academic preparation for college, and primarily on ELA and mathematics. States tend to treat career readiness as synonymous with college readiness. When states include career readiness in their accountability systems, they tend to use narrow indicators, or they only address non-cognitive skills at the school level. The construct of readiness encompasses more than academics, and despite some progress by the states, none have fully addressed college and career readiness in their assessments and accountability systems, and information about students’ readiness remains incomplete.

75 See https://www.kuder.com/
76 See https://www.naviance.com/
77 See http://www.wowi.com/
References


National Assessment Governing Board

Executive Committee

Thursday, November 15, 2018
4:30 – 6:00 pm

AGENDA

4:30 – 4:35 pm  Welcome and Agenda Overview
    Bev Perdue, Chair
    Action: Executive Committee Membership

4:35 – 5:20 pm  CLOSED: Long-Term Trend Budget and Schedule
    Lisa Stooksberry, Deputy Executive Director
    Peggy Carr, Associate Commissioner, NCES
    Attachment A

5:20 – 6:00 pm  CLOSED: Preview of the NAEP Assessment Schedule
    and Budget Briefing
    Peggy Carr
    Attachment B
Long Term Trend Budget and Schedule:
Appropriations Update for Fiscal Year 2019

On September 28, 2018 the President signed the Fiscal Year 2019 appropriations bill for the Department of Education which provides steady funding for the Governing Board ($7.745 million) and a $2 million increase for the NAEP program ($151 million total) for October 1, 2018 – September 30, 2019.

The appropriations law includes the following requirement for the Governing Board regarding the use of these additional NAEP funds:

“The conferees direct the National Assessment Governing Board to brief the Committees on Appropriations of the House of Representatives and the Senate within 60 days from the date of enactment of this Act on the resources required to administer a long-term trend assessment by 2021.”

In this closed session, the Executive Committee will be briefed by NCES on the budget estimates and operational steps to conduct the Long-Term Trend assessment (LTT) earlier than its next scheduled administration in 2024. The Executive Committee will consider the Board’s response to the appropriations committees, which is due on Tuesday, November 27th.
On Thursday, November 15, the Executive Committee will receive a preview of the NAEP Assessment Schedule and Budget Briefing to be conducted in plenary session on Friday, November 16. The following materials are provided as background for both budget sessions.

National Assessment Governing Board

Closed Session: NAEP Assessment Schedule and Budget

Setting the NAEP Assessment Schedule is one of the Governing Board’s most important statutory responsibilities. Historically, the Governing Board has amended the NAEP Assessment Schedule to reflect legislative changes to NAEP’s authorization, new opportunities, and evolving expectations in what students should know and be able to do. According to the Governing Board's General Policy on Conducting and Reporting NAEP, the Board “periodically establishes a dependable, publicly announced assessment schedule of at least ten years in scope. The schedule specifies the subject or topic (e.g., High School Transcript Study), grades, ages, assessment year, and sampling levels (i.e., national, state, TUDA) for each assessment.” The current Schedule of Assessments (attached) was approved in November 2015 and extends through 2024.

The Board’s Strategic Vision includes a priority to “Develop policy approaches to revise the NAEP assessment subjects and schedule based on the nation’s evolving needs, the Board’s priorities, and NAEP funding” (SV #9). To begin pursuing this strategic priority, Governing Board members engaged in small group and plenary discussions on this topic during several Board meetings over the past year. These discussions culminated in the adoption of a Resolution on Board Priorities for the NAEP Assessment Schedule (attached) at the March 2018 Board meeting.

During the May 2018 Board meeting, Governing Board members engaged in small group discussions to consider various approaches for implementing the assessment schedule priorities of utility, frequency, and efficiency. In its discussion, the Board generally agreed that it is desirable to increase state and TUDA administrations for Reading, Mathematics, Science, and Writing in particular.

During the August 2018 Board meeting, there was a plenary presentation and discussion of several potential approaches to increase the efficiency of the U.S. History, Civics, Geography, and Economics assessments. Relative benefits and costs of each approach were discussed, including potential implications for trends, achievement levels, and reporting. Several Board members noted that a consolidated social studies framework that would result in an overall social studies score and achievement levels may be too broad to be meaningful, helpful, or actionable. There was some support for separate assessments and trends in Civics and U.S. History to be maintained (even if the administrations are coordinated to produce results about interrelationships) and possibly prioritized over Geography and Economics.

Based on the Board’s previous discussions and operational considerations provided by NCES, Governing Board and NCES staff will share the rationale and projected costs associated with a proposed draft schedule through 2030.
National Assessment Governing Board Resolution on Priorities for the NAEP Assessment Schedule

Whereas, The Nation’s Report Card—also known as the National Assessment of Educational Progress (NAEP)—is mandated by Congress to conduct a national assessment and report data on student academic achievement and trends in public and private elementary schools and secondary schools (P.L. 107-279);

Whereas, the NAEP Authorization Act requires that NAEP be administered in public and private schools in reading and mathematics at least every 2 years in grades 4 and 8 and every 4 years in grade 12 and conduct the Long-Term Trend assessment in reading and mathematics for ages 9, 13, and 17;

Whereas, the NAEP Authorization Act specifies that beyond the requirements listed above, to the extent time and resources allow, NAEP shall assess and report achievement trends in additional subjects in grades 4, 8, and 12;

Whereas, the Every Student Succeeds Act mandates that states participate in the biennial reading and mathematics NAEP assessments in grades 4 and 8;

Whereas, Congress supported the establishment and expansion of the NAEP Trial Urban District Assessment (TUDA) to provide NAEP results for select large urban districts;

Whereas, NAEP provides national, state, and local policymakers and practitioners with consistent, external, independent measures of student achievement through which results across education systems can be compared at points in time and over time;

Whereas, the National Assessment Governing Board and the National Center of Education Statistics (NCES) continuously work to enhance NAEP’s form (e.g. transitioning to digital-based assessments) and content (e.g. the Technology and Engineering Literacy assessment) to reflect the modern expectations of what students know and can do;

Whereas, Congress authorized the National Assessment Governing Board to determine the NAEP subjects to be assessed;

Whereas, it is the National Assessment Governing Board’s policy, in consultation with NCES, to periodically establish a dependable, publicly announced NAEP Schedule of Assessments spanning at least ten years, and specifying the subjects, grades, ages, assessment years, sampling levels (e.g., national, state, TUDA), and introduction of new and revised frameworks for each assessment;

Whereas, on November 18, 2016 the National Assessment Governing Board unanimously adopted its Strategic Vision which included a priority to “Develop policy approaches to revise the NAEP assessment subjects and schedule based on the nation’s evolving needs, the Board priorities, and NAEP funding“;
Therefore, as the National Assessment Governing Board anticipates extending the NAEP Schedule of Assessments into the future, it will uphold all of the aforementioned requirements and make decisions informed by each of the following priorities to ensure NAEP results are impactful and policy-relevant:

- **Utility** – include more voluntary state and Trial Urban District Assessments and continue to align the schedule of NAEP administrations with international assessments in the same subjects to enable actionable comparisons of districts, states, and other nations;

- **Frequency** – commit to assess subjects other than reading and mathematics at least every 4 years to provide additional measures of student academic progress at regular intervals; and

- **Efficiency** – find cost-effective ways to administer NAEP while to the degree possible maintaining a breadth of subjects on the schedule in order to continue reporting progress in student achievement;

Furthermore, the National Assessment Governing Board recognizes that any change to the NAEP Schedule of Assessments requires consideration of the fiscal, technical, and operational implications.
Schedule Information by Subject

Reading
- NAEP legislation specifies every 2 years at grades 4 and 8 for nation and states; NCLB/ESSA requires states to partake
- NAEP legislation specifies every 4 years at grade 12 for nation
- Administration has included voluntary TUDAs for grades 4 and 8 since 2002
- Administered at national level only for grade 12, and for 11-13 states voluntarily participated in 2009 and 2013
- Grade 12 assessment used to estimate % of students academically prepared for college
- Current trend lines begin in 1992
- Administration coincides with PIRLS (grade 4) once every 10 years

Math
- NAEP legislation specifies every 2 years at grades 4 and 8 for nation and states; NCLB/ESSA requires states to partake
- NAEP legislation specifies every 4 years at grade 12 for nation
- Administration has included voluntary TUDAs for grades 4 and 8 since 2003
- Administered at national level only for grade 12, and for 11-13 states voluntarily participated in 2009 and 2013
- Grade 12 assessment used to estimate % of students academically prepared for college
- Current trend lines begin in 1990 for grades 4 and 8; 2005 for grade 12
- Administration coincides with every administration of TIMSS (4 year cycle)

Science
- Has been administered approximately every 4 years at all 3 grades
- Administered to the nation, states, and (usually) voluntary TUDAs for grades 4 and 8
- Administered at national level only for grade 12
- Current trend lines begin in 2009
- Since 2011, administration has coincided with every administration of TIMSS

Writing
- Has been administered approximately every 4 years at grades 8 and 12; much less frequently at grade 4
- Under current framework (beginning with 2011 administration), has been administered to the nation only
- Previous framework included administration to states and voluntary TUDAs in 1998 (states only), 2002, 2007
History
- Has been administered at the national level approximately every 4 years at grade 8; less frequently at grades 4 and 12

Civics
- Has been administered at the national level approximately every 4 years at grade 8; less frequently at grades 4 and 12

Geography
- Has been administered at the national level approximately every 4 years at grade 8; less frequently at grades 4 and 12

Technology and Engineering Literacy (TEL)
- Has been administered at national level for grade 8 only in 2014 and 2018
- Framework covers all 3 grades

Economics
- Framework covers grade 12 only
- Has been administered at national level in 2006 and 2012

Arts
- Framework covers all 3 grades but administered at national level for grade 8 only
- Framework includes 4 areas (Dance, Music, Visual Arts, and Theatre) but only Music and Visual Arts have been included in operational assessment
- New framework is needed for transition to DBA; not feasible to complete in time for 2024 administration

Foreign Language
- Framework to measure Spanish language proficiency adopted in 2000
- Pilot test conducted in 2003 but assessment never administered operationally

High School Transcript Study
- Supplemental data collection to grade 12 Math and Science administrations
- NCES has been working to determine the feasibility of conducting this study for grade 8 and at the state level

Long-Term Trend (LTT)
- Legislation notes continuing for Reading and Math, but no periodicity specified
- Periodicity has varied but generally has been at least every 4 years until 2012
# National Assessment of Educational Progress

## Schedule of Assessments

Approved November 21, 2015

The *National Assessment of Educational Progress (NAEP) Authorization Act* established the National Assessment Governing Board to set policy for NAEP, including determining the schedule of assessments. (P.L. 107-279)

<table>
<thead>
<tr>
<th>Year</th>
<th>Subject</th>
<th>National Grades Assessed</th>
<th>State Grades Assessed</th>
<th>TUDA Grades Assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>U.S. History*</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Civics*</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Geography*</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>TECHNOLOGY AND ENGINEERING LITERACY</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>2015</td>
<td>Reading*</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td></td>
<td>Mathematics*</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td></td>
<td>Science**</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td>2016</td>
<td>Arts*</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>2017</td>
<td>Reading</td>
<td>4, 8</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td></td>
<td>Mathematics</td>
<td>4, 8</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td></td>
<td>Writing</td>
<td>4, 8</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td>2018</td>
<td>U.S. History</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Civics</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Geography</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Technology and Engineering Literacy</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>2019</td>
<td>Reading</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td></td>
<td>Mathematics</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td></td>
<td>Science</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td></td>
<td>High School Transcript Study</td>
<td>4, 8</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td>2020</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2021</td>
<td>Reading</td>
<td>4, 8</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td></td>
<td>Mathematics</td>
<td>4, 8</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td></td>
<td>Writing</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td>2022</td>
<td>U.S. HISTORY</td>
<td>8, 12</td>
<td>8, 12</td>
<td>8, 12</td>
</tr>
<tr>
<td></td>
<td>CIVICS</td>
<td>8, 12</td>
<td>8, 12</td>
<td>8, 12</td>
</tr>
<tr>
<td></td>
<td>GEOGRAPHY</td>
<td>8, 12</td>
<td>8, 12</td>
<td>8, 12</td>
</tr>
<tr>
<td></td>
<td>Economics</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Technology and Engineering Literacy</td>
<td>8, 12</td>
<td>8, 12</td>
<td>8, 12</td>
</tr>
<tr>
<td>2023</td>
<td>Reading</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td></td>
<td>Mathematics</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td></td>
<td>Science</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td></td>
<td>High School Transcript Study</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2024</td>
<td>ARTS</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>FOREIGN LANGUAGE</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Long-term Trend</td>
<td>~</td>
<td>~</td>
<td>~</td>
</tr>
</tbody>
</table>

**NOTES:**

- *Assessments not administered by computer. Beginning in 2017 all operational assessments will be digitally based.
- **Science in 2015 consisted of paper-and-pencil and digital-based components.
- ~Long-term Trend (LTT) assessments sample students at ages 9, 13, and 17 and are conducted in reading and mathematics.
- Subjects in **BOLD ALL CAPS** indicate the year in which a new framework is implemented or assessment year for which the Governing Board will decide whether a new or updated framework is needed.
## AGENDA

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
<th>Presenter(s)</th>
<th>Attachment</th>
</tr>
</thead>
<tbody>
<tr>
<td>10:30 – 11:05 am</td>
<td>Welcome and Opening Remarks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11:05 – 11:20 am</td>
<td>CLOSED SESSION: Vocabulary Assessment in NAEP Reading: Latest NCES Research and Development</td>
<td>Eunice Greer, NCES</td>
<td>Attachment A</td>
</tr>
<tr>
<td>11:20 – 11:50 am</td>
<td>Committee Recommendation for the NAEP Reading Framework (SV#5)</td>
<td>Carol Jago, ADC Chair</td>
<td>Attachment B</td>
</tr>
<tr>
<td>12:20 – 12:30 pm</td>
<td>Briefing: Assessment Systems in Other Countries (SV#8)</td>
<td>Art Thacker, Human Resources Research Organization (HumRRO)</td>
<td>Attachment D</td>
</tr>
<tr>
<td></td>
<td>Next Steps &amp; Questions on Information Items</td>
<td>Carol Jago, ADC Chair</td>
<td></td>
</tr>
</tbody>
</table>

### Information Items

- Implementation of ADC Strategic Vision Activities
- Item Review Schedule

Attachments:
- Attachment A
- Attachment B
- Attachment C
- Attachment D
- Attachment E
- Attachment F
NAEP VOCABULARY ASSESSMENT: UPDATE AND OVERVIEW

The current Reading Framework calls for the assessment of “meaning vocabulary”. Meaning vocabulary is defined as “…the application of one’s understanding of word meanings to passage comprehension.” (2009 NAEP Reading Framework, p. 32). Initially, most NAEP reading blocks, which comprised a passage and 10 – 13 items, included 2 – 3 meaning vocabulary items. In order to have sufficient numbers of items to yield reliable scale scores, “reading half-blocks” were created. These blocks included 2 passages, selected for their rich vocabulary. These half-block passages were half the length of normal NAEP passages and each half-block passage included 5 – 6 vocabulary items. These vocabulary blocks were spiraled with the normal, full-length passage blocks. Vocabulary scale scores were reported, through 2015, but the high correlations between meaning vocabulary coupled with the restricted range of performance on the items, called into question the usefulness of the measure.

In November of 2016, NCES convened a panel of experts in vocabulary learning and assessment. This team began to work together as the NAEP Vocabulary Panel. The primary charge to the Vocabulary Panel was to recommend ways to enhance the operationalization of Meaning Vocabulary assessment in NAEP Reading as a means of improving the Meaning Vocabulary measure. The overarching principle guiding the work of the Vocabulary Panel was the same as that guiding the current Meaning Vocabulary measure—i.e., Meaning Vocabulary items are meant to assess students’ use of words to build text understanding, not simply knowledge of isolated words. (2009 NAEP Reading Framework, p. 32). The panel members worked together over a period of six months to: (1) integrate current research on vocabulary knowledge and assessment with the existing guidelines for Meaning Vocabulary and (2) create prototypes for new items.

This presentation will serve to bring the Assessment Development Committee up to date on the status of NAEP’s assessment of vocabulary within the Reading assessment. The presentation will address four topics:

- The history of NAEP’s efforts to assess meaning vocabulary
- The relevant research that informed the Vocabulary Panel’s recommendations for new items and formats
- The Vocabulary Panel’s recommendations for new Meaning Vocabulary item types, including examples of new, piloted items from 4th and 8th grade
- Recommendations for Expanding Meaning Vocabulary as it is addressed in the NAEP Reading Framework
The Assessment Development Committee (ADC) hosted a panel discussion in March 2018 and collected several expert papers as part of its Review of the NAEP Reading Framework. In April and August 2018, the ADC met via teleconference to discuss and refine the Committee’s recommendation on the NAEP Reading Framework. The Committee’s recommendation is intended to set the scope of the framework update and indicate the Board’s priorities. If the Committee recommends a major update, the recommendation will also serve as a charge for the Framework Visioning Panel that will be convened by the Governing Board to develop detailed recommendations for the framework update. In accordance with the Governing Board Framework Development Policy, the Visioning Panel shall include a wide array of leading voices in reading.

A draft of the Committee’s recommendation follows. At the November 2018 Board meeting, ADC Chair Carol Jago will invite additional discussion and comments in order to finalize the recommendation for full Board action in March 2019.
The National Assessment Governing Board Charge to the Visioning and Development Panels For the 2025 National Assessment of Educational Progress (NAEP) Reading Framework

Whereas. The Nation’s Report Card—also known as the National Assessment of Educational Progress (NAEP)—is mandated by Congress to conduct national assessments and report data on student academic achievement and trends in public and private elementary schools and secondary schools, and is prohibited from using any assessment to “evaluate individual students or teachers” or “to establish, require, or influence the standards, assessments, curriculum, … or instructional practices of states or local education agencies” (Public Law 107-279);

Whereas. Congress specifically assigned the National Assessment Governing Board responsibilities to “develop assessment objectives consistent with the requirements of this [law] and test specifications that produce an assessment that is valid and reliable, and are based on relevant widely accepted professional standards”;

Whereas. the Governing Board’s Strategic Vision adopted in November 2016 established that the Board will, “develop new approaches to update NAEP subject area frameworks to support the Board's responsibility to measure evolving expectations for students, while maintaining rigorous methods that support reporting student achievement trends”;

Whereas. the Governing Board established in its Framework Development Policy that the Board shall conduct “a comprehensive, inclusive, and deliberative process” to determine and update the content and format of all NAEP assessments;

Whereas. in accordance with the Governing Board’s Framework Development Policy, the Board’s Assessment Development Committee conducted a review of the current NAEP Reading Framework, which included seven papers from leading reading educators;

Whereas. based on the review of the NAEP Reading Framework conducted by the Assessment Development Committee, the Committee concludes that a substantial framework update is required to address digital platforms and new research, and recommends that the Board update the NAEP Reading Framework last updated in 2004 “to be informed by a broad, balanced, and inclusive set of factors” balancing “current curricula and instruction, research regarding cognitive development and instruction, and the nation’s future needs and desirable levels of achievement,” in accordance with the Framework Development Policy;

Therefore,

- The National Assessment Governing Board staff, with appropriate contractor support and oversight by the Governing Board’s Assessment Development Committee, shall conduct a framework update by establishing a Visioning Panel with a subset of members continuing as the Development Panel, in accordance with the Governing Board Framework Development Policy;

- All processes and procedures identified in the Governing Board Framework Development Policy shall be followed;
• The Visioning and Development Panels will recommend necessary changes in the NAEP Reading Framework at grades 4, 8, and 12 that maximize the value of NAEP to the nation; and the Panels are also tasked with considering opportunities to extend the depth of measurement and reporting given the affordances of digital based assessment;

• The update process shall result in three documents: a recommended framework, assessment and item specifications, and recommendations for contextual variables that relate to student achievement in reading;

• At the conclusion of the NAEP Reading Framework update process, the National Assessment Governing Board shall review recommendations from the Visioning and Development Panels, and take final action on recommended updates to the reading framework, assessment specifications, and subject-specific contextual variables; and

• The framework update adopted by the Board will guide development of the 2025 NAEP Reading Assessment.
Quarterly Report

2025 NAEP MATHEMATICS FRAMEWORK UPDATE

WestEd
November 2018
Contract #91995918C0001
Quarterly Progress Report

Project Overview

In September 2018, the Governing Board awarded a contract to WestEd to conduct an update of the NAEP Mathematics and Reading Assessment Frameworks, Test and Item Specifications, and Contextual Variables. The goal of the first year of the project is to update the mathematics framework documents through the work of a 30-person Visioning Panel, a 15-person Development Panel, and an 8-person Technical Advisory Committee (TAC). This will be accomplished through an initial Visioning Panel meeting, three subsequent Development Panel meetings, conducting outreach efforts to gather public comment on draft versions of the documents, and production of a final updated Assessment Framework, Test and Item Specifications, and Contextual Variables for Mathematics to submit to the Governing Board by July 2019. The NAEP Reading Assessment Framework, Test and Item Specifications, and Contextual Variables will be updated in Year 2 of the project.

The work is to be conducted using a combination of external experts and mathematics specialists within WestEd. WestEd’s considerable experience with NAEP comes from having led two previous Framework projects: 1) the update of the 2009 NAEP Science Assessment Framework, and 2) the development of an assessment framework for a new 2014 NAEP assessment in Technology and Engineering Literacy (TEL). To complete this work, WestEd is also partnering with the Council of Chief State School Officers (CCSSO), who will assist in compiling resources for the Framework panels and in securing feedback on the updated framework, test and item specifications, and contextual variables. Input into the framework document update will also come from project collaborators: the National Council of Teachers of Mathematics (NCTM), the National Council of Supervisors of Mathematics (NCSM), TODOS: Mathematics for All (TODOS), and the California Mathematics Project (CMP).

Project Team

Suzanne Wilson, Professor in the Department of Curriculum and Instruction, University of Connecticut will serve as the Mathematics Panel Chair, leading the Visioning and Development Panel activities. The Project Management Team consists of Steve Schneider, Mark Loveland, Ann Edwards, Matt Gaertner, and Kellie Kim. As project director, Steve Schneider will provide day-to-day leadership, guidance, and liaising with the Governing Board. Dr. Schneider has over 40 years of science, mathematics, and technology education experience and led WestEd’s two previous Framework development projects. Project co-director, Mark Loveland, and Mathematics Content Expert, Ann Edwards, have oversight for all programmatic activities. Dr. Loveland was project coordinator for the TEL Framework development project, and Dr. Edwards has over 20 years of experience in research, curriculum development, and
professional development in mathematics education spanning elementary to post-secondary levels. Measurement Expert, Dr. Gaertner will support TAC Chair, Guillermo Solano-Flores, Professor of Education at the Stanford University Graduate School of Education. Dr. Kim will serve as Process Manager, facilitating panel meetings and documenting all project activities. In addition, the broader project team includes two mathematics subject matter experts, a project coordinator, and research assistants.

**Project Plan**

The project plan involves WestEd project management and coordination of panel and TAC activities to update the NAEP Mathematics Assessment Framework, Test and Item Specifications, and Contextual Variables. The framework update deliberations and drafting will be carried out by the Framework Visioning and Development Panels. Comprised of 30 individuals representing various stakeholder groups, the Framework Visioning Panel will formulate guidelines for developing a recommended framework, based on the state of the field. Fifteen members of the Visioning Panel will constitute the Framework Development Panel. The Development Panel is charged with developing the drafts of the three project documents and engaging in the detailed deliberations to determine how to reflect the Visioning Panel guidelines in an updated framework. Dates for the Visioning Panel and Development Panel meetings have been finalized; the Visioning Panel meeting is scheduled for November 7-8, 2018, and the three Development Panel meetings are scheduled for December 5-6, 2018, January 9-10, 2019, and February 12-13, 2019. All meetings will take place in the Washington, DC area.

Preparatory work for the Framework Panel activities has been extensive. WestEd has prepared a Project Plan, which describes the process and schedule for updating the framework documents, and a project Design Document, which serves as the blueprint for the project processes, describing outcomes and metrics, and as the touchstone for quality assurance monitoring.

Using processes outlined in the Design Document, WestEd has worked in consultation with Governing Board staff and Governing Board members to identify a final list of 30 members of the Visioning and Development Panels. To inform the development of a final list of prospective members, the project Management Team, with input from mathematics subject matter experts, constructed a matrix of potential panelists, arranged according to key stakeholder group representation and issues/areas of expertise. Consideration was also given to organizational representation, variables of interest for panelist expertise or background including grade band, teaching experience, geographic region, locale (i.e. urban, suburban, rural), and gender. Additionally, WestEd has identified a final list of 8 technical experts specializing in educational measurement to comprise the TAC. The TAC will respond to technical issues raised during panel deliberations.

The work of the panels and TAC will be informed by an Issues Paper and a compilation of resources. The Issues Paper is intended to serve as a springboard for discussion by the Framework panels and will address issues that are likely to be engaged in the update of the NAEP Mathematics Framework. The Issues Paper and resource compilation will draw from the current NAEP Mathematics framework, specifications, and contextual variables documents, national and international standards, state frameworks and standards, extant assessments, reports, research on mathematics education and
assessment, and other resources. These resources highlight reports, high-level presentations, and associated academic research papers, including an alignment study of NAEP, PARCC\(^1\), Smarter Balanced, and several state assessment frameworks. Using the Issues Paper and documents identified in the compilation of resources, project staff have prepared a presentation and discussion process to use in the initial orientation of the panels to their responsibilities in updating the NAEP Mathematics Framework.

Upon completion of the Framework panel meetings, WestEd will facilitate the actual updating of the Mathematics framework documents, producing draft versions of the Assessment Framework, Test and Item Specifications, and Contextual Variables. Outreach will be conducted primarily by WestEd and CCSSO, with assistance from collaborating organizations. Feedback on the draft documents will come from member organizations represented on the two panels, other organizations, and the public. Organizations may choose to convene meetings, gather feedback via web-based survey, or have members contact a web site. In all instances, the groups will follow procedures for securing input and ensuring representation of diverse views. WestEd staff will tabulate feedback, make recommendations for revisions addressing the feedback, and coordinate the development of final versions of the framework documents, to be submitted to the Governing Board.

## Milestones

The major milestones of the project are summarized below.

<table>
<thead>
<tr>
<th>Milestone</th>
<th>Estimated Timing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Kickoff Meeting</td>
<td>09/20/18</td>
</tr>
<tr>
<td>Project Plan Development</td>
<td>09/27/18 – 10/07/18</td>
</tr>
<tr>
<td>Design Document Development</td>
<td>10/11/18 – 10/24/18</td>
</tr>
<tr>
<td>Identification of Visioning and Development Panelists and TAC Members</td>
<td>09/20/18 – 10/10/18</td>
</tr>
<tr>
<td>Issues Paper and Resource Compilation Development</td>
<td>09/20/18 – 11/1/18</td>
</tr>
<tr>
<td>Visioning Panel Meeting</td>
<td>11/7-8/18</td>
</tr>
<tr>
<td>Development Panel Meetings</td>
<td>12/5-6/18</td>
</tr>
<tr>
<td></td>
<td>01/9-10/19</td>
</tr>
<tr>
<td></td>
<td>02/12-13/19</td>
</tr>
<tr>
<td>Convene TAC</td>
<td>Three weeks after each panel meeting and prior to submission of draft framework documents</td>
</tr>
<tr>
<td>Draft Versions of Framework Documents</td>
<td>04/19/19</td>
</tr>
<tr>
<td>Gather Public Comment</td>
<td>04/22/19 – 06/14/19</td>
</tr>
<tr>
<td>Develop Final Versions of Framework Documents</td>
<td>04/19/18 – 06/30/19</td>
</tr>
<tr>
<td>Submit Final Framework Documents to Governing Board</td>
<td>07/01/19</td>
</tr>
<tr>
<td>Submit Final Process Report</td>
<td>09/16/19</td>
</tr>
</tbody>
</table>

\(^1\) The Partnership for Assessment of Readiness for College and Careers
ASSESSMENT SYSTEMS IN OTHER COUNTRIES

As the Governing Board Technical Services contractor, the Human Resources Research Organization (HumRRO) has developed several resources to inform innovations in how NAEP framework updates are developed and implemented. HumRRO is currently developing a technical memorandum profiling assessment systems in other countries, which will support upcoming framework projects with ideas to consider for frameworks, framework processes, contextual data, and reporting. Art Thacker of HumRRO will brief the ADC on major takeaways from this review.
IMPLEMENTATION OF ADC STRATEGIC VISION ACTIVITIES

The ADC develops recommendations for what NAEP should assess and exercises final authority over all NAEP items. Each framework development and update process uses broadly representative panels of stakeholders to develop detailed recommendations as a draft NAEP assessment framework. Each Board-adopted NAEP framework describes what students should know and be able to do in a subject area and what will be tested on NAEP.

Several activities in the Governing Board Strategic Vision call for ADC’s leadership. These projects involve informing educators, updating policies, and exploring new approaches to framework updating, as well as projects to review and update frameworks as needed. A working draft of ADC’s project plans is attached, along with a summary of common elements for each framework project. For a detailed timeline of the NAEP Mathematics Framework Update process, see Attachment C.

Ongoing Committee Discussions

Recent ADC discussions have raised several issues for ongoing discussion as the Committee leads Strategic Vision activities and prepares content recommendations for Board deliberation and action:

- The optimal role of NAEP for each content area.
- How Board and Committee priorities should be reflected in upcoming framework updates.
- Expected gains and losses for each NAEP assessment decision.
- Extent to which current frameworks are flexible enough to adapt as needed.
- The level of specificity in assessment results that is most useful to policymakers, researchers, and educators.
- How future NAEP items will be a resource for the field.
- How to establish and maintain partnerships that highlight actionable aspects of results, e.g., teacher access to released NAEP items and contextual information.
- How to develop viable options for new configurations of NAEP assessment content in ways that balance expertise, outreach, research, and trends in curricular standards.
- How to incorporate how other countries think about changing what they assess.
- Whether to more deeply assess an existing content area or add new content areas.
- Whether streamlining of NAEP frameworks is an appropriate goal.
- How to be intentional about content overlap between different assessments, while fulfilling statutory requirements, e.g., biennial reading and mathematics assessment.
## Working Draft* Plan: All ADC Strategic Vision (SV) Activities

### Updates Since August 2018 Are Highlighted

<table>
<thead>
<tr>
<th>Activity</th>
<th>Start</th>
<th>Finish</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identify NAEP Resources &amp; Information for Educators (SV #3 Expanding NAEP Resources and SV #6 Contextual Variables)</td>
<td>May 2017</td>
<td>Nov 2021</td>
<td>ADC discussed NAEP Questions Tool and contextual variables in 2017. Suggestions for new or refined NAEP resources can be shared with R&amp;D for Board outreach. To be determined: when/how to develop ADC recommendations.</td>
</tr>
<tr>
<td>Review &amp; Update Mathematics Framework for 2025 Assessment</td>
<td>Aug 2017</td>
<td>Mar 2025</td>
<td>State math standards review began in August 2017. Results were shared in May 2018 ADC Framework Review, which also engaged external expert commentary. ADC prepared a framework recommendation for Board action, and it was unanimously adopted in August 2018. The framework contractor(^2) for the Math Framework Update project was secured in Summer 2018. The Board will review a draft framework when public comment is being collected in Spring 2019. Board action is slated for Summer/Fall 2019, allowing NCES to conduct development leading to a 2025 administration of the updated assessment.</td>
</tr>
<tr>
<td>Review &amp; Update Reading Framework for 2025 Assessment</td>
<td>Oct 2017</td>
<td>Mar 2025</td>
<td>ADC Framework Review was held in March 2018 to inform development of recommendations for a Fall 2019 framework update project launch. In August 2018, the ADC prepared a draft framework recommendation for Board action. Discussion will continue at the November 2018 Board meeting, so the recommendation can be finalized for Board action in March 2019.</td>
</tr>
<tr>
<td>Explore New Approaches to Framework Update Processes (also SV #8 International Assessments)</td>
<td>Nov 2017</td>
<td>Aug 2023</td>
<td>The Board’s Technical Services contractor is developing several resources to assist in exploring innovations in how NAEP assessment updates are implemented. Framework Update Projects will review other countries’ assessment programs to inform frameworks, framework processes, contextual data, and reporting.</td>
</tr>
</tbody>
</table>

* All timelines are estimated. This draft will be updated based on Board policy decisions. All activities address Strategic Vision Priority #5 Updating Frameworks, unless otherwise noted.
1 Timeline includes administering the assessment.
2 The mathematics framework project will be implemented by the same contractor as the reading framework project, on staggered schedules so that most of the mathematics project is completed by the time the reading project begins.
**WORKING DRAFT PLAN: ALL ADC STRATEGIC VISION (SV) ACTIVITIES**

*Updates since August 2018 are Highlighted*

<table>
<thead>
<tr>
<th>Activity</th>
<th>Start</th>
<th>Finish</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Update Item Development Policy</td>
<td>Aug 2018</td>
<td>Aug 2019</td>
<td>The ADC began discussing goals for the policy revision in August 2018. In Fall 2018, an expert panel will be convened to gather insights from the field regarding best practices in assessment development.</td>
</tr>
<tr>
<td>Review &amp; Update Civics, Geography, and U.S. History Frameworks (Depends on NAEP Schedule)</td>
<td>Mar 2018</td>
<td>May 2020</td>
<td>Discussion of outreach began in March 2018, with suggestions to develop options for the ADC to consider. In August 2018, ADC review of the current NAEP item pools indicated that framework revisions did not need to be fast-tracked. Framework reviews will begin in 2019, which will include external expert commentary.</td>
</tr>
<tr>
<td>Review &amp; Update Economics Framework (Depends on NAEP Schedule)</td>
<td>Mar 2020</td>
<td>Aug 2021</td>
<td>Depending on ADC recommendations and Board Assessment Schedule decisions, Economics may or may not be a standalone project.</td>
</tr>
<tr>
<td>Review &amp; Update Science and Technology &amp; Engineering Literacy (TEL) Frameworks (Depends on NAEP Schedule)</td>
<td>Sep 2020</td>
<td>Nov 2022</td>
<td>Discussion of outreach began in March 2018, Tentative next steps: learn more about standards in NGSS non-adopter states and learn whether stakeholders view that some or all of the TEL subarea on Technology &amp; Society addresses student achievement goals in Civics, Geography, U.S. History, or Economics.</td>
</tr>
<tr>
<td>Develop Content Descriptions for the Long-Term Trend (LTT) Mathematics and Reading Assessments (SV #7 Long-Term Trend)</td>
<td>TBD</td>
<td>TBD</td>
<td>March 2018 Executive Committee deliberations on LTT called for ADC to develop content descriptions of the assessments to support LTT item development, as well as updates to the Governing Board LTT policy and improved explanations of LTT assessment goals. ADC requested these descriptions also illuminate knowledge and skills of lower performing students, if possible. NCES has already developed a list of measurement objectives for LTT Mathematics, and similar work may be possible for Reading. Board staff is using these inputs to begin development of the LTT content descriptions.</td>
</tr>
</tbody>
</table>
Common Elements of Each Framework Update Project

Based on the revised Framework Development Policy, several milestones address all NAEP assessment framework projects. Framework update projects engage stakeholders and content experts to identify needed revisions, via subject-specific factors including:

- Evolution of discipline and implications for NAEP frameworks
- Relevance to students’ postsecondary endeavors
- Student achievement trends in terms of contextual factors
- Digital-based assessment issues
- International content and measurement trends

<table>
<thead>
<tr>
<th>MILESTONES: ALL FRAMEWORK PROJECTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADC Discussion with External Experts in the Subject Area(s)</td>
</tr>
<tr>
<td>ADC Recommendation for Updating Assessment</td>
</tr>
<tr>
<td>Board Action on Charge</td>
</tr>
<tr>
<td>Framework Contractor Selection</td>
</tr>
<tr>
<td>Trend Scan &amp; Resource Compilation</td>
</tr>
<tr>
<td>Panel Meetings (3 to 6)</td>
</tr>
<tr>
<td>Full Board Review &amp; Public Comment</td>
</tr>
<tr>
<td>Framework Draft Finalized</td>
</tr>
<tr>
<td>ADC Final Review of Framework</td>
</tr>
<tr>
<td>Board Action</td>
</tr>
<tr>
<td>Assessment Administered</td>
</tr>
</tbody>
</table>

As a first step, the ADC conducts a framework review, where content experts are invited to a Committee session to provide reflections on the state of the discipline and the extent to which the relevant NAEP framework should be updated. Studies and additional outreach is pursued, as needed, to inform the ADC’s recommendation about the type of framework update that is required. Next, the ADC brings its recommendation to the full Board for approval. In the case of an anticipated framework update, the recommendation includes a charge to stakeholders who will serve on the panels convened to draft recommendations for the ADC’s consideration.

After Board discussion of the ADC recommendation, the Board will take action on the charge. Concurrently, Board staff will identify a contractor to execute the framework update process.

The framework contractor will launch the project by identifying individuals to serve on the framework panels and by compiling and developing resources to support the meetings of these stakeholders. A subset of these resources will include the Governing Board’s charge.
to the framework panels as well as documents used to inform the Board’s development of the charge. The first meeting of stakeholders will be for the Visioning Panel to discuss the major issues to be addressed in the framework. A subset of the Visioning Panel will continue on as the Development Panel to develop an updated framework. This panel will also develop the recommended updates to the Test and Item Specifications, as well as the Contextual Variables.

The ADC monitors the framework contractor’s work via regular project updates. A draft of the panels’ recommended framework will be shared for full Board review and public comment, as well as review by the Board’s Committee on Standards, Design and Methodology. This feedback will allow the Development Panel to address concerns and finalize the draft framework, specifications, and contextual variables for the ADC’s final review and Board action. The adopted framework, specifications, and contextual variables are given to NCES to begin assessment development, piloting, and finally administration of the operational assessment based on the new framework.
## Assessment Development Committee
### Item Review Schedule
#### June 2018 – May 2019
##### Updated October 12, 2018

<table>
<thead>
<tr>
<th>Review Package to Board</th>
<th>Board Comments to NCES</th>
<th>Survey/ Cognitive</th>
<th>Review Task</th>
<th>Approx. Number Items</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>12/3/2018</td>
<td>12/21/18</td>
<td>Cognitive</td>
<td>2022 TEL (8 &amp; 12) Pilot Concept Sketches</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>2/13/18</td>
<td>3/8/18</td>
<td>Survey</td>
<td>2023 Reading (4, 8, 12) Existing Pool Review</td>
<td>TBD</td>
<td></td>
</tr>
<tr>
<td>2/13/18</td>
<td>3/8/18</td>
<td>Survey</td>
<td>2023 Mathematics (4, 8, 12) Existing Pool Review</td>
<td>TBD</td>
<td></td>
</tr>
<tr>
<td>2/13/18</td>
<td>3/8/18</td>
<td>Survey</td>
<td>2023 Science (4, 8, 12) Existing Pool Review</td>
<td>TBD</td>
<td></td>
</tr>
<tr>
<td>2/21/18</td>
<td>3/8/18</td>
<td>Cognitive</td>
<td>2023 Reading (4, 8, 12) Pilot (SBT) Concept Sketch</td>
<td>8-12</td>
<td></td>
</tr>
<tr>
<td>2/21/18</td>
<td>3/8/18</td>
<td>Cognitive</td>
<td>2023 Science (4, 8, 12) Pilot (SBT) Concept Sketch</td>
<td>5-10</td>
<td></td>
</tr>
<tr>
<td>4/17/2019</td>
<td>5/10/2019</td>
<td>Cognitive</td>
<td>2022 Civics (8, 12) Pilot</td>
<td>184-204</td>
<td></td>
</tr>
<tr>
<td>4/17/2019</td>
<td>5/10/2019</td>
<td>Cognitive</td>
<td>2022 Geography (8, 12) Pilot</td>
<td>168-188</td>
<td></td>
</tr>
</tbody>
</table>

**NOTE:** “SBT” indicates Scenario-Based Task
“DI” indicates Discrete Item.
## National Assessment Governing Board  
### Committee on Standards, Design and Methodology  
#### Friday, November 16, 2018  
**10:30 am – 12:30 pm**

### AGENDA

<table>
<thead>
<tr>
<th>Time</th>
<th>Activity</th>
<th>Presenter(s)</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>10:30 – 10:35 am</td>
<td>Welcome and Review of Agenda</td>
<td><em>Andrew Ho, COSDAM Chair</em></td>
<td></td>
</tr>
<tr>
<td>10:35 – 10:40 am</td>
<td><strong>ACTION:</strong> Policy Statement on Developing Student Achievement Levels for NAEP</td>
<td><em>Andrew Ho</em></td>
<td>See Achievement Levels Policy Tab</td>
</tr>
</tbody>
</table>
| 10:40 – 11:20 am | Update on Implementing the Board’s Response to the Evaluation of NAEP Achievement Levels | *Sharyn Rosenberg, Assistant Director for Psychometrics*  
*Andrew Ho* | *Attachment A*                                                            |
| 11:20 – 11:25 am | Questions on Information Items (see below)                               |                                                                              |                                |
| 11:30 am – 12:30 pm | Joint Session with Reporting & Dissemination Committee: Communication and Interpretation of NAEP Achievement Levels (SV #3) | *Rebecca Gagnon, R&D Chair*  
*Andrew Ho* | *Attachment B*                                                            |
|                 | **Information Item**                                                     |                                                                              |                                |
|                 | Update on Implementing the Strategic Vision (SV #2-10)                   |                                                                              | *Attachment C*                 |
Implementing the Governing Board’s Response to the 2016 Evaluation of NAEP Achievement Levels

The final report of the most recent evaluation of NAEP achievement levels was released on November 17, 2016; a free PDF of the full report can be downloaded at: https://www.nap.edu/catalog/23409/evaluation-of-the-achievement-levels-for-mathematics-and-reading-on-the-national-assessment-of-educational-progress. The Governing Board received a briefing from staff at the National Academies of Sciences, Engineering, and Medicine and members of the interdisciplinary review committee during the November 2016 Board meeting. As required by law, the Governing Board adopted a formal response to the evaluation (see attached) that was sent to the Secretary of Education, the Committee on Education and the Workforce of the House of Representatives, and the Committee on Health, Education, Labor, and Pensions of the Senate on December 20, 2016.

Over the past couple of years, COSDAM has had several conversations about how to implement various aspects of the Board’s response to the evaluation. The table below summarizes the current status, recent work, and planned next steps for each of the recommendations. During the upcoming November Board meeting, COSDAM members will engage in discussions about how to proceed with various activities, including potential collaborations with NCES and the Reporting and Dissemination Committee.

| Recommendations and Board Response                                                                 | Current Status and Recent Work                                                                                                                                                                                                 | Planned Next Steps                                                                                                                                                                                                                                                                 |
|-----------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------ Adam londenucmnt 1: Evaluating the alignment of NAEP achievement level descriptors (ALDs)  
*The Governing Board intends to issue a procurement for conducting studies to achieve this goal*  
*The updated Board policy on NAEP achievement levels will address the larger issue of specifying a process and timeline for conducting regular recurring reviews of the ALDs in all subjects and grades (in conjunction with Recommendation #3)* | After the proposed revised policy is adopted, the procurement work can begin, including any additional COSDAM discussion of the study design  
There are funds in the Governing Board’s Fiscal Year 2019 budget to issue a procurement for this work  
In August 2018, COSDAM had a preliminary discussion about this work and agreed that reporting would begin with the 2021 math and reading results  
In July 2018, HumRRO convened an expert panel to discuss the approach for review and revision of ALDs and developing reporting ALDs under the proposed revised policy (for more detail, see the attached minutes)  
The proposed revised policy that COSDAM has been working on since March 2017 addresses ALDs, process for reporting ALDs, periodic review of all ALDs  
The updated Board policy on NAEP achievement levels will address the larger issue of specifying a process and timeline for conducting regular recurring reviews of the ALDs in all subjects and grades (in conjunction with Recommendation #3) |
<table>
<thead>
<tr>
<th>Recommendations and Board Response</th>
<th>Current Status and Recent Work</th>
<th>Planned Next Steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recommendation #2: Determination of the trial status of the NAEP achievement levels</td>
<td>The evaluation stated that only the first recommendation (followed by an additional evaluation) was necessary for removing the trial designation, but this decision is at the discretion of the NCES Commissioner</td>
<td>Implement the Board’s response to this evaluation. NCES will determine the appropriate time for conducting an additional evaluation. At the completion of a new evaluation, the then-current NCES Commissioner will make a determination about whether the trial designation should be removed.</td>
</tr>
<tr>
<td><strong>The NCES Commissioner is responsible for determining whether the trial designation is removed</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recommendation #3: Regular recurring reviews of the ALDs</td>
<td>This is addressed by Principle 4 (Periodic Review of Achievement Levels) in the proposed revised policy</td>
<td>Adopt the proposed revised policy. Review and revision of math and reading ALDs will occur first, but other subjects will follow shortly thereafter after the policy is adopted.</td>
</tr>
<tr>
<td><strong>The revised policy will include a statement of periodicity for conducting regular recurring reviews of the ALDs, with updates as needed</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recommendation #4: Relationships between NAEP achievement levels and external measures</td>
<td>Governing Board and NCES staff presented an initial idea of how to synthesize existing linking study findings at the May 2017 COSDAM meeting. Additional information about current and possible future linking studies was shared with COSDAM in August 2017 and March 2018.</td>
<td>Additional work is needed to figure out a comprehensive plan for conducting and reporting this research in collaboration with NCES.</td>
</tr>
<tr>
<td><strong>The Governing Board and NCES have some additional linkages planned and underway</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>The Governing Board anticipates that additional linkages with external measures will help connect the NAEP achievement levels and scale scores to other meaningful real-world indicators of current and future performance</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recommendations and Board Response</td>
<td>Current Status and Recent Work</td>
<td>Planned Next Steps</td>
</tr>
<tr>
<td>------------------------------------</td>
<td>------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Recommendation #5: Interpretations and uses of NAEP achievement levels</td>
<td>To inform the statement of intended uses of NAEP, HumRRO has been working on a synthesis of existing uses of NAEP, which can also inform future data collection efforts</td>
<td>Develop statement of intended uses and interpretations of the NAEP achievement levels</td>
</tr>
<tr>
<td><strong>The Governing Board will issue a procurement to conduct research to better understand how various audiences have used and interpreted NAEP results (including achievement levels)</strong></td>
<td>There have been several COSDAM discussions about developing a statement of intended uses for NAEP, including for the achievement levels (most recently at the August 2018 COSDAM meeting; see attached)</td>
<td>Perform additional research to better understand how various audiences are interpreting achievement levels and how communications can be improved</td>
</tr>
<tr>
<td><strong>The Governing Board will work collaboratively with NCES to provide further guidance and outreach about appropriate and inappropriate uses of NAEP achievement levels</strong></td>
<td>There was a discussion about considerations for developing a validity argument for the NAEP achievement levels at the March 2018 COSDAM meeting (see attached technical memo)</td>
<td>Develop interpretative guides to be linked to the Nations Report Card</td>
</tr>
<tr>
<td></td>
<td>The proposed revised policy refers to interpretative guide to accompany NAEP releases; there will be a discussion of how to approach this guide (and defining uses of NAEP achievement levels) at the upcoming November 2018 joint session with R&amp;D</td>
<td>Develop a validity argument for the NAEP achievement levels based on the intended uses and interpretations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Perform additional outreach</td>
</tr>
<tr>
<td>Recommendation #6: Guidance for inferences made with achievement levels versus scale scores</td>
<td>This work is contingent upon having a statement of intended uses of NAEP and should be incorporated into the interpretative guide, in addition to other communication materials</td>
<td>Develop a plan to implement this recommendation, in conjunction with the R&amp;D Committee and NCES</td>
</tr>
<tr>
<td><strong>The Governing Board will continue to work with NCES and follow current research to provide guidance about inferences that are best made with achievement levels and those best made with scale score statistics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recommendations and Board Response</td>
<td>Current Status and Recent Work</td>
<td>Planned Next Steps</td>
</tr>
<tr>
<td>---------------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| Recommendation #7: Regular cycle for considering desirability of conducting a new standard setting | *The Governing Board will update its policy on setting achievement levels for NAEP to be more explicit about conditions that require a new standard setting* | Adopt revised policy statement  
Consider whether any current standards should be reviewed |
National Assessment Governing Board’s Response to the National Academies of Sciences, Engineering, and Medicine 2016 Evaluation of NAEP Achievement Levels

Legislative Authority

Pursuant to the National Assessment of Educational Progress (NAEP) legislation (Public Law 107-279), the National Assessment Governing Board (hereafter the Governing Board) is pleased to have this opportunity to apprise the Secretary of Education and the Congress of the Governing Board response to the recommendations of the National Academies of Sciences, Engineering, and Medicine evaluation of the NAEP achievement levels for mathematics and reading (Edley & Koenig, 2016).

The cited legislation charges the Governing Board with the authority and responsibility to “develop appropriate student achievement levels for each grade or age in each subject area to be tested.” The legislation also states that “such levels shall be determined by... a national consensus approach; used on a trial basis until the Commissioner for Education Statistics determines, as a result of an evaluation under subsection (f), that such levels are reasonable, valid, and informative to the public; ... [and] shall be updated as appropriate by the National Assessment Governing Board in consultation with the Commissioner for Education Statistics” (Public Law 107-279).

Background

NAEP is the largest nationally representative and continuing assessment of what our nation’s elementary and secondary students know and can do. Since 1969, NAEP has been the country’s foremost resource for measuring student progress and identifying differences in student achievement across student subgroups. In a time of changing state standards and assessments, NAEP serves as a trusted resource for parents, teachers, principals, policymakers, and researchers to compare student achievement across states and select large urban districts. NAEP results allow the nation to understand where more work must be done to improve learning among all students.

For 25 years, the NAEP achievement levels (Basic, Proficient, and Advanced) have been a signature feature of NAEP results. While scale scores provide information about student achievement over time and across student groups, achievement levels reflect the extent to which student performance is “good enough,” in each subject and grade, relative to aspirational goals.
Since the Governing Board began setting standards in the early 1990s, achievement levels have become a standard part of score reporting for many other assessment programs in the US and abroad.

Governing Board Response

Overview

The Governing Board appreciates the thorough, deliberative process undertaken over the past two years by the National Academies of Science, Engineering, and Medicine and the expert members of the Committee on the Evaluation of NAEP Achievement Levels for Mathematics and Reading. The Governing Board is pleased that the report concludes that the achievement levels are a meaningful and important part of NAEP reporting. The report states that, “during their 24 years [the achievement levels] have acquired meaning for NAEP’s various audiences and stakeholders; they serve as stable benchmarks for monitoring achievement trends, and they are widely used to inform public discourse and policy decisions. Users regard them as a regular, permanent feature of the NAEP reports” (Edley & Koenig, 2016; page Sum-8). The Governing Board has reviewed the seven recommendations presented in the report and finds them reasonable and thoughtful. The report will inform the Board’s future efforts to set achievement levels and communicate the meaning of NAEP Basic, Proficient, and Advanced. The recommendations intersect with two Governing Board documents, the Strategic Vision and the achievement levels policy, described here.

On November 18, 2016, the Governing Board adopted a Strategic Vision (https://www.nagb.org/content/nagb/assets/documents/newsroom/press-releases/2016/nagb-strategic-vision.pdf) to guide the work of the Board through 2020, with an emphasis on innovating to enhance NAEP’s form and content and expanding NAEP’s dissemination and use. The Strategic Vision answers the question, “How can NAEP provide information about how our students are doing in the most innovative, informative, and impactful ways?” The Governing Board is pleased that several of the report recommendations are consistent with the Board’s own vision. The Governing Board is committed to measuring the progress of our nation’s students toward their acquisition of academic knowledge, skills, and abilities relevant to this contemporary era.

The Governing Board’s approach to setting achievement levels is articulated in a policy statement, “Developing Student Performance Levels for the National Assessment of Educational Progress” (https://www.nagb.org/content/nagb/assets/documents/policies/developing-student-performance.pdf). The policy was first adopted in 1990 and was subsequently revised in 1995,
with minor wording changes made in 2007. The report motivates the revision of this policy, to add clarity and intentionality to the setting and communication of NAEP achievement levels.

The seven recommendations and the Governing Board response comprise a significant research and outreach trajectory that the Governing Board can pursue over several years in conjunction with key partners. The Governing Board will implement these responses within resource constraints and in conjunction with the priorities of the Strategic Vision.

**Evaluating the Alignment of NAEP Achievement Level Descriptors**

**Recommendation #1:** Alignment among the frameworks, the item pools, the achievement-level descriptors, and the cut scores is fundamental to the validity of inferences about student achievement. In 2009, alignment was evaluated for all grades in reading and for grade 12 in mathematics, and changes were made to the achievement-level descriptors, as needed. Similar research is needed to evaluate alignment for the grade 4 and grade 8 mathematics assessments and to revise them as needed to ensure that they represent the knowledge and skills of students at each achievement level. Moreover, additional work to verify alignment for grade 4 reading and grade 12 mathematics is needed.

The report’s primary recommendation is to evaluate the alignment, and revise if needed, the achievement level descriptors for NAEP mathematics and reading assessments in grades 4, 8, and 12. The Governing Board intends to issue a procurement for conducting studies to achieve this goal. The Governing Board has periodically conducted studies to evaluate whether the achievement level descriptors in a given subject should be revised, based on their alignment with the NAEP framework, item pool, and cut scores. The Governing Board agrees that this is a good time to ensure that current NAEP mathematics and reading achievement level descriptors align with the knowledge and skills of students in each achievement level category. In conjunction with the response to Recommendation #3, the updated Board policy on NAEP achievement levels will address the larger issue of specifying a process and timeline for conducting regular recurring reviews of the achievement level descriptions in all subjects and grades.

The Governing Board agrees strongly with the recommendation that, while evaluating alignment of achievement level descriptors is timely, it is not necessary to consider changing the cut scores or beginning a new trend line at this time. The NAEP assessments are transitioning from paper-based to digital assessments in 2017, and current efforts are focused on ensuring comparability between 2015 and 2017 scores. The Governing Board articulated this in the 2015 Resolution on Maintaining NAEP Trends with the Transition to Digital-Based Assessments ([https://www.nagb.org/content/nagb/assets/documents/policies/resolution-on-trend-and-dba.pdf](https://www.nagb.org/content/nagb/assets/documents/policies/resolution-on-trend-and-dba.pdf)).

**Recommendation #2:** Once satisfactory alignment among the frameworks, the item pools, the achievement-level descriptors, and the cut scores in NAEP mathematics and reading has been
demonstrated, their designation as trial should be discontinued. This work should be completed and the results evaluated as stipulated by law: (20 U.S. Code 9622: National Assessment of Educational Progress: https://www.law.cornell.edu/uscode/text/20/9622 [September 2016]).

Ultimately, the Commissioner of Education Statistics is responsible for determining whether the “trial” designation is removed. The Governing Board is committed to providing the Commissioner with the information needed to make this determination in an expedient manner.

**Regular Recurring Reviews of the Achievement Level Descriptors**

**Recommendation #3:** To maintain the validity and usefulness of achievement levels, there should be regular recurring reviews of the achievement-level descriptors, with updates as needed, to ensure they reflect both the frameworks and the incorporation of those frameworks in NAEP assessments.

The Board’s current policy on NAEP achievement levels contains several principles and guidelines for setting achievement levels but does not address issues related to the continued use or reporting of achievement levels many years after they were established. The revised policy will seek to address this gap by including a statement of periodicity for conducting regular recurring reviews of the achievement level descriptors, with updates as needed, as called for in this recommendation. The Governing Board agrees that it is important to articulate a process and timeline for conducting regular reviews of the achievement level descriptors rather than performing such reviews on an ad hoc basis.

**Relationships Between NAEP Achievement Levels and External Measures**

**Recommendation #4:** Research is needed on the relationships between the NAEP achievement levels and concurrent or future performance on measures external to NAEP. Like the research that led to setting scale scores that represent academic preparedness for college, new research should focus on other measures of future performance, such as being on track for a college-ready high school diploma for 8th-grade students and readiness for middle school for 4th-grade students.

In addition to the extensive work that the Governing Board has conducted at grade 12 to relate NAEP mathematics and reading results to academic preparedness for college, the Governing Board has begun research at grade 8 with statistical linking studies of NAEP mathematics and reading and the ACT Explore assessments in those subjects. This work was published while the evaluation was in process and was not included in the Committee’s deliberations. Additional studies in NAEP mathematics and reading at grades 4 and 8 are beginning under contract to the National Center for Education Statistics (NCES). The Governing Board’s Strategic Vision includes an explicit goal to increase opportunities for connecting NAEP to other national and
international assessments and data. Just as the Board’s previous research related grade 12 NAEP results in mathematics and reading to students’ academic preparedness for college, the Governing Board anticipates that additional linkages with external measures will help connect the NAEP achievement levels and scale scores to other meaningful real-world indicators of current and future performance.

**Interpretations and Uses of NAEP Achievement Levels**

*Recommendation #5: Research is needed to articulate the intended interpretations and uses of the achievement levels and collect validity evidence to support these interpretations and uses. In addition, research to identify the actual interpretations and uses commonly made by NAEP’s various audiences and evaluate the validity of each of them. This information should be communicated to users with clear guidance on substantiated and unsubstantiated interpretations.*

The Governing Board’s Strategic Vision emphasizes improving the use and dissemination of NAEP results, and the Board’s work in this area will include achievement levels. The Governing Board recognizes that clarity and meaning of NAEP achievement levels (and scale scores) are of utmost importance. The Governing Board will issue a procurement to conduct research to better understand how various audiences have used and interpreted NAEP results (including achievement levels). The Governing Board will work collaboratively with NCES to provide further guidance and outreach about appropriate and inappropriate uses of NAEP achievement levels.

**Guidance for Inferences Made with Achievement Levels versus Scale Scores**

*Recommendation #6: Guidance is needed to help users determine inferences that are best made with achievement levels and those best made with scale score statistics. Such guidance should be incorporated in every report that includes achievement levels.*

The Governing Board understands that improper uses of achievement level statistics are widespread in the public domain and extend far beyond the use of NAEP data. Reports by the Governing Board and NCES have modeled appropriate use of NAEP data and will continue to do so. This recommendation is also consistent with the goal of the Strategic Vision to improve the dissemination and use of NAEP results. The Governing Board will continue to work with NCES and follow current research to provide guidance about inferences that are best made with achievement levels and those best made with scale score statistics.
Regular Cycle for Considering Desirability of Conducting a New Standard Setting

Recommendation #7: NAEP should implement a regular cycle for considering the desirability of conducting a new standard setting. Factors to consider include, but are not limited to: substantive changes in the constructs, item types, or frameworks; innovations in the modality for administering assessments; advances in standard setting methodologies; and changes in the policy environment for using NAEP results. These factors should be weighed against the downsides of interrupting the trend data and information.

When the Board’s achievement levels policy was first created and revised in the 1990s, the Board was setting standards in each subject and grade for the first time and had not yet considered the need or timeline for re-setting standards. To address this recommendation, the Governing Board will update the policy to be more explicit about conditions that require a new standard setting.

Board’s Commitment

The Governing Board remains committed to its congressional mandate to set “appropriate student achievement levels” for the National Assessment of Educational Progress. The Board appreciates the report’s affirmation that NAEP achievement levels have been set thoughtfully and carefully, consistent with professional guidelines for standard setting, and based on extensive technical advice from respected psychometricians and measurement specialists. The Board also takes seriously the charge to develop the current achievement levels through a national consensus approach, involving large numbers of knowledgeable teachers, curriculum specialists, business leaders, and members of the general public throughout the process. This is only fitting given the Governing Board’s own congressionally mandated membership that explicitly includes representatives from these stakeholder groups.

The Governing Board remains committed to improving the process of setting and communicating achievement levels. The Governing Board is grateful for the report recommendations that will advance these aims.

Reference

Notes of the Expert Panel Meeting on NAEP Achievement Level Descriptions
July 12-13, 2018
National Assessment Governing Board

Contents

Background................................................................................................................................... 1
Presentations ................................................................................................................................ 2
Discussion..................................................................................................................................... 8
  Proposed Revisions to the Board Policy on Achievement Level Setting ........................................9
  Purposes of Reporting ALDs .................................................................................................... 9
  Audiences for Reporting ALDs ................................................................................................10
  Various Item Mapping Methods for ALD Development ...........................................................10
  “Top Down” and “Bottom Up” ALD Development ....................................................................12
  Panelists and Procedures for Validating ALDs and Developing Reporting ALDs .....................15
  Special Considerations for Math and Reading Reporting ALDs ..............................................16
  Steps to Validate Alignment of Current ALDs and Develop Reporting ALDs .........................17
  Validation/Vetting of Reporting ALDs ......................................................................................17
Summary and Reflections........................................................................................................... 18
References.................................................................................................................................. 19

Appendix A: Meeting Agenda, Attendees, and List of Read-Ahead Materials .........................20
  July 12 –13, 2018 │ Agenda ............................................................................................... 21
  Attendees ............................................................................................................................ 22
  Read-ahead Materials......................................................................................................... 23

Appendix B: Presentations.......................................................................................................... 24
  Overview of the National Assessment Governing Board and NAEP ALDs (Sharyn Rosenberg) .........................................................................................................................25
  Reporting Achievement Level Descriptors (Hillary Michaels) ..................................................31
  Validating ALDs: What Have the States Done? (Marianne Perie) .............................................34
  Anchor Studies for NAEP Achievement Levels (Susan Loomis) .............................................35
  The Basis of Scale Anchoring in Item Mapping: Some Issues of Concern (Andrew Kolstad) 37
List of Figures

Figure 1. Depiction of ALD development: Current Governing Board policy and procedures ........2

Figure 2. Depiction of ALD development: Proposed revised Governing Board policy and procedures .................................................................4

Figure 3. Hypothetical outcomes associated with RP80. (Figure reproduced from Kolstad (2018).) .................................................................................................7

Figure 4. Top-down development of content ALDs from higher level policy ALDs. ...............13

Figure 5. Bottom-up development of hypothetical reporting ALDs from items that map onto a given NAEP achievement level. .................................................................14
Notes of the Expert Panel Meeting on Achievement Level Descriptions

July 12–13, 2018
National Assessment Governing Board

At the request of the National Assessment Governing Board, HumRRO organized and facilitated a meeting with a select group of leading experts in assessment and achievement level setting. The purpose of this meeting was to elicit input about proposed changes to the development and use of NAEP achievement level descriptions (ALDs) in general and about how to approach the first recommendation from the recent evaluation of NAEP achievement levels (National Academies of Sciences, Engineering, and Medicine, 2017), which is focused on reviewing and revising the NAEP ALDs for mathematics and reading. The meeting was timed to share high level outcomes with the Committee on Standards, Design and Methodology (COSDAM) as they consider policy changes and next steps for addressing the first recommendation from the evaluation.

The purpose of this document is to summarize the themes and comments from the rich discussions at the meeting.

Background

We were fortunate to assemble an exceptional panel of experts (hereafter referred to as “the Experts”): **Dr. Susan Davis-Becker**, ACS Ventures; **Dr. Karla Egan**, EdMetric; **Dr. Steve Ferrara**, Measured Progress; **Dr. Ed Haertel**, Stanford University; **Dr. Andrew Kolstad**, P20 Strategies, **Dr. Susan Loomis**, Consultant; **Dr. Barbara Plake**, University of Nebraska-Lincoln, and **Dr. Lauress Wise**, HumRRO. The full list of meeting attendees is included in Appendix A.

Four papers on aspects of ALDs were commissioned and provided as read-ahead materials in advance of the panel meeting:

- **Reporting Achievement Level Descriptors for the National Assessment of Educational Progress;** Dr. Hillary Michaels, HumRRO; Dr. Karla Egan, EdMetric; Dr. Art Thacker, HumRRO; and Dr. Sheila Schultz, HumRRO
- **Validating Achievement Level Descriptors;** Dr. Marianne Perie, University of Kansas
- **Anchor Studies for Analysis of NAEP Achievement Levels;** Dr. Susan Loomis, Consultant
- **The Basis of Scale Anchoring in Item Mapping: Some Issues of Concern;** Dr. Andrew Kolstad, P20 Strategies

The meeting was held on July 12–13, 2018 in Alexandria, Virginia. In advance of the meeting, the Experts received an agenda and the following read-ahead materials: (a) the National Academies of Sciences, Engineering, and Medicine’s *Evaluation of the Achievement Levels for Mathematics and Reading on the National Assessment of Educational Progress*, (b) the Governing Board’s formal response to the evaluation of NAEP Achievement Levels, (c) the four papers listed in the above bulleted list, and (d) an example of a previous NAEP anchor study. Appendix A contains the agenda and list of read-ahead materials.

Dr. Sunny Becker, HumRRO, welcomed the Experts and led the attendees through introductions. She also reviewed the agenda and stated the goals for the meeting.
The morning of the first day was devoted to a series of presentations. Copies of all presentation slides are in Appendix B. The remainder of the meeting comprised group discussions of several topics, including (a) use and range of reporting ALDs, (b) methodology, (c) panelists and procedures, (d) special considerations for mathematics and reading ALDs, (e) special considerations for reporting ALDs when setting new achievement levels, (f) validation/vetting of reporting ALDs, and (g) recommendations for special studies. Because this meeting was held so close in time to the COSDAM meeting it was meant to inform, there was insufficient time to prepare and vet these meeting notes to submit to COSDAM. In lieu of these notes, Dr. Sharyn Rosenberg offered a summary of important themes for the Governing Board and COSDAM to consider regarding the development and revision of reporting ALDs; the Experts agreed to this summary.

Presentations

To set the context for the Experts, Dr. Sharyn Rosenberg, Assistant Director for Psychometrics on the Governing Board staff, provided an overview of NAEP achievement level setting, including current and proposed roles of the ALDs throughout the process. Figure 1 depicts the current process,

![Figure 1. Depiction of ALD development: Current Governing Board policy and procedures](image)

Table 1 describes the types of ALDs, and uses of each, according to the current policy and procedures.

<table>
<thead>
<tr>
<th>Policy definitions</th>
<th>The current policy on Developing Student Performance Levels for NAEP defines three NAEP achievement levels: Basic, Proficient, and Advanced. These policy definitions apply to all subjects and grade levels for which NAEP achievement levels are set.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preliminary ALDs</td>
<td>The current policy refers to preliminary ALDs, which are developed by the framework committee. These preliminary ALDs typically have separate statements for each content area and grade level, and are intended to inform item development (as described by the Board policy on Item Development and Review). The statements are written in terms of what students should know and be able to do.</td>
</tr>
<tr>
<td>Final ALDs</td>
<td>The current policy refers to final ALDs but is ambiguous about when the ALDs are revised and “locked down” from further changes. Since 1998, the preliminary ALDs have been reviewed and revised by a panel of content experts prior to beginning the achievement level setting activities. The rationale for finalizing the ALDs in advance of (rather than during) the achievement level setting meetings is to allow for thoughtful review and vetting (including public comment) on the final ALDs aimed to assure appropriate alignment to the policy definitions across achievement levels within each grade and across grades within each achievement level. The final ALDs typically contain summary statements for the</td>
</tr>
</tbody>
</table>
subject overall and may also contain additional details by content area. The statements are written in terms of what students should know and be able to do.

<table>
<thead>
<tr>
<th>Threshold ALDs (if applicable)</th>
<th>If descriptions of performance right at the cut scores are needed for the standard setting methodology (e.g., Bookmark), then threshold (or borderline) ALDs are developed by achievement level setting panelists. Panelists are told that the threshold ALDs are for their own use only and will not be reported with the NAEP results.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALDs in NAEP reports</td>
<td>The ALDs currently included in NAEP reports are generally the same as the final ALDs used in the achievement level setting. An exception to this practice is when there were framework changes that required revisions to the ALDs (e.g., 2009 reading) or changes to cut scores that necessitated development of new ALDs (e.g., 1996 science). In several other cases, anchoring studies were conducted to evaluate the validity of the existing final ALDs but did not result in changes to the ALDs for reporting (see Loomis paper in Appendix A).</td>
</tr>
</tbody>
</table>

Dr. Rosenberg noted that the Governing Board is in the process of revising its policy on developing student achievement levels for NAEP. COSDAM developed a draft revised policy, but this document was not shared in the read-ahead materials because the initial full Board discussion was planned for the August Board meeting, shortly after the panel meeting took place.

In her presentation, Dr. Rosenberg shared excerpts from the proposed revised policy that were directly related to ALDs to inform the discussion:

- There are a few minor edits to the policy definitions for clarity, including: adding “NAEP” in front of Basic, Proficient, and Advanced to better differentiate the NAEP achievement levels from other uses of these terms; and removing the term “grade” to avoid confusion with “grade-level” performance.

- ALDs for a specific grade and subject (e.g., NAEP grade 4 mathematics) are collectively referred to as “content ALDs” to differentiate from the policy definitions (e.g., NAEP Proficient) that apply to performance on all NAEP assessments.

- Content ALDs are developed initially as part of the framework development process and may be revised to serve other purposes such as guiding an achievement level setting. The content ALDs that guide achievement level setting activities shall be written in terms of what students should know and be able to do.

- There will be no content ALDs developed for performance below the NAEP Basic level (consistent with the current policy).

- When content ALDs are reported with results (also referred to as reporting ALDs), they shall be written to incorporate empirical data from student performance. They shall describe what students do know and can do rather than what they should know and should be able to do (this represents a major change from the current policy and practices).
- There is a new principle on periodic review of achievement levels (and ALDs), to address one of the recommendations from the recent evaluation of NAEP achievement levels.

Figure 2 depicts the proposed revised process.

![Figure 2. Depiction of ALD development: Proposed revised Governing Board policy and procedures](image)

Table 2 describes the proposed types and uses of ALDs.

### Table 2. Types and Uses of ALDs: Proposed Revised Governing Board Policy and Procedures

<table>
<thead>
<tr>
<th>Policy definitions</th>
<th>The proposed revised policy on Developing Student Achievement Levels for NAEP defines three NAEP achievement levels: NAEP Basic, NAEP Proficient, and NAEP Advanced. These policy definitions apply to all main NAEP assessments.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content ALDs</td>
<td>ALDs in Framework (two types - for item development and achievement level setting) Under the revised policy and procedures for framework development, the framework committee will develop content ALDs both by content area (to inform item development) and overall (for use in the achievement level setting activities). These ALDs will continue to be written in terms of what students should know and be able to do. If there is a specific need to revise the overall ALDs in advance of an achievement level setting, then a separate activity will be undertaken to do so, but this is not intended to be necessary in most cases.</td>
</tr>
<tr>
<td>Threshold ALDs (if applicable)</td>
<td>If descriptions of performance right at the cut scores are needed for the standard setting methodology (e.g., Bookmark), then threshold (or borderline) ALDs will continue to be developed by achievement level setting panelists. Panelists are told that the threshold ALDs are for their own use only and will not be reported with the NAEP results.</td>
</tr>
<tr>
<td>Reporting ALDs</td>
<td>The proposed revised policy calls for conducting a study following an achievement level setting to revise the content ALDs for the purpose of reporting, using empirical data of student performance. The reporting ALDs will be written in terms of what students do know and can do.</td>
</tr>
</tbody>
</table>

Dr. Rosenberg explained that she wanted to understand the technical feasibility and any challenges associated with developing reporting ALDs for NAEP based on empirical data.

She shared the first recommendation from the recent evaluation of NAEP achievement levels in mathematics and reading:
“Alignment among the frameworks, the item pools, the achievement-level descriptors, and the cut scores is fundamental to the validity of inferences about student achievement. In 2009, alignment was evaluated for all grades in reading and for grade 12 in mathematics, and changes were made to the achievement level descriptors, as needed. Similar research is needed to evaluate alignment for the grade 4 and grade 8 mathematics assessments and to revise them as needed to ensure that they represent the knowledge and skills of students at each achievement level. Moreover, additional work to verify alignment for grade 4 reading and grade 12 mathematics is needed” (National Academies of Sciences, Engineering, and Medicine, 2017).

In its formal response to this recommendation, the Governing Board pledged to conduct studies to review and revise the ALDs in math and reading at grades 4, 8, and 12. Dr. Rosenberg noted that one of the primary goals from this meeting is to seek input on considerations for conducting these studies.

Next, one author of each of the four commissioned papers summarized the paper and answered related questions.

**Dr. Hillary Michaels** shared information from a recent paper that she, Karla Egan of EdMetric, and other HumRRO colleagues prepared on developing reporting ALDs (Michaels et al., 2018). She emphasized the primary purpose of reporting ALDs is to communicate clearly to stakeholders the knowledge, skills, and abilities that students can demonstrate at each achievement level. Although reporting ALDs have been used for years, she noted the major challenge with reporting ALDs is that they typically are not developed specifically for who or how they will be used. She described four types of ALDs—of which reporting ALDs are one type—as well as the intended purpose and primary audience for each. Although ALDs are developed and used by various stakeholders, policy ALDs are often developed for policy makers and politicians; range ALDs are developed for teachers, content and curriculum experts and item writers; and reporting ALDs are developed to inform the general public and media representatives. Dr. Michaels also reviewed an example of policy, range, and reporting ALDs and discussed the type of information presented in each as well as the differences in level of specificity among them. She ended her presentation by having participants consider (a) what is the proper language to use for reporting ALDs (statements that use would vs. should vs. could), (b) whether content ALDs should be revised following standards setting (thereby creating final ALDs), (c) the extent to which information about use should be included in the dissemination of reporting ALDs, and (d) the importance of gathering validity evidence when developing and using reporting ALDs.

**Dr. Marianne Perie** joined the meeting remotely to provide a summary of approaches used by states to validate ALDs on their assessments (Perie, 2018). She discussed four goals for ALDs and the evidence states collect to demonstrate validity of their achievement levels: (a) ALDs should be fully aligned to the assessment; (b) ALDs should provide an accurate representation of student knowledge and skills; (c) ALDs should follow a clear progression across levels and grades; and (d) ALDs should be clearly written and easily understandable by a larger audience. Dr. Perie described five approaches states use to validate ALDs: (a) alignment, (b) item mapping, (c) item descriptors, (d) student mapping, and (e) survey. Alignment methods focus on using an evidence-centered design approach to match the rigor of the content in the ALD to the rigor of the framework and items. In item mapping, experts can review an item and match the content of that item to an appropriate ALD or they use item statistics to map the item to an appropriate ALD. The item descriptors approach is similar to item mapping, with the difference being the use of groups of items rather than individual items. Dr. Perie explained that a student mapping approach is not appropriate for NAEP because there are no individual scores. Surveys
are used to look at readability, clarity, appropriateness, and utility of items. Similar to what states often do, Dr. Perie suggested using multiple approaches for NAEP, including item descriptors and surveys of the intended target audience.

**Dr. Susan Loomis** provided an historical context by describing previous NAEP anchor studies (Loomis, 2018). The NAS Report recommended conducting studies that evaluate consistency among NAEP frameworks, item pools, ALDs, and cut scores. Although each study is designed to answer specific research questions, the general purpose of an anchor study is to determine the extent to which student performance, within achievement level ranges, demonstrates the knowledge, skills, and abilities described in the achievement level descriptors. Anchor studies are part of the validity evidence supporting the NAEP cut scores and ALDs.

Dr. Loomis highlighted several studies included in her paper, all of which used some type of item mapping technique:

- **Science 1996 Grades 4, 8, and 12** when data from standard setting panels were used as the basis for the Governing Board’s adjustments to cut scores as recommended by standard setting panelists. Two panels developed ALDs from the items order at response probability (RP) 0.67. The Governing Board adopted the ALDs developed by one of the panels.
- **Science 2009 Grades 4, 8, and 12** when all Grade 4 cut scores and Advanced cut scores for Grades 8 and 12 were changed.
- **Reading 2009 Grades 4, 8, and 12** when a new framework was implemented, and the 1992 cut scores and score scale were unchanged. Panelists used anchor descriptions to develop ALDs for reporting the 2009 Reading results relative to the 1992 Reading cut scores.
- **Geography 2002 Grades 4, 8, and 12** using all student data and conditional probabilities for performance scores. This two-year study was conducted in phases. The first phase investigated whether student performance was represented by the ALDs developed in 1994. The second phase convened two independent panels of geography experts. One panel had experts familiar with NAEP while the other panel included experts who were not. Both panels wrote descriptions of items from the Grade 8 2001 assessment that had been anchored to the scale for the study. The results indicated a high level of consistency between the groups.
- **Mathematics 2003 Anchor Study Grades 4 and 8.** The study looked at the extent to which the changes in NAEP item pools over the years since the framework was first implemented had impacted the alignment of items with the 1992 ALDs. The studies found the ALDs remained generally consistent over time in describing what students should know and do even though item types had changed, particularly in geometry.

Dr. Loomis noted considerations for panelist recruitment and materials required to conduct anchor studies to review and revise the mathematics and reading ALDs. Based on previous experience, she suggested the panelists and facilitators should be qualified in their content area and be familiar with NAEP. She also recommended there be two panels of about six panelists each. Panelists should review items from more than one assessment cycle and would review all item types. The ALDs should be considered final until the framework on which they are based changes. In her experience, these guidelines have resulted in clear and generalizable results.

The following additional methodological considerations were discussed: (a) applying a consistent RP criterion, such as 0.67, in both standard setting and anchor studies; (b) including a measure of discrimination; (c) not using a correction for guessing for anchor studies unless used for standard setting; (d) including a focus on items that do not anchor; (e) conducting two-
way comparisons of descriptions and alignment ratings; and (f) including evaluations by panelists at key points in the process.

**Dr. Andrew Kolstad** discussed concerns with the convention of using a response probability (RP) value of 67 for standards setting (Kolstad, 2018). His concerns stem from the potentially misguided idea of “item mastery.” He described a standards setting in which panelists determine a point in an item booklet ordered by item difficulty where students described by an ALD no longer get items correct. An assumption underlying this method is that the student would get all of the prior (easier) items correct, and all the later (harder) items incorrect. However, if the student missed an earlier item, to get the same score, they would need to get one of the later items correct. Dr. Kolstad argued that this method would create a mismatch between the description of the student indicated by the ALD classification and the KSAs the student actually possessed.

One of the problems with using an RP value of greater than 50 is that it increases the number of false negative classifications of students. A greater number of students who can actually do the things indicated in the ALD will be classified into a lower category (false negative) than the number of students who cannot do those things will be classified into a higher category (false positive). Dr. Kolstad illustrated this phenomenon using a graphic of the items a particular student got correct versus those incorrect using an RP value of 80. In the graphic, he labels both the false negative and false positive classifications. It is clear from the illustration that error is skewed toward the negative when RP is greater than 50. The illustration from Dr. Kolstad’s paper is reproduced below in Figure 3.

![Figure 3. Hypothetical outcomes associated with RP80. (Figure reproduced from Kolstad (2018).)](image)

Dr. Kolstad also provided examples of how basing ALDs on higher RP values can be very misleading, quoting a New York Times article indicating that half of Americans had limited proficiency with English. The article provided an illustrative item from the proficient category of an English assessment and indicated that half of Americans did not get questions similar to the illustrative question correct. In fact, 72% of Americans got the illustrative item correct, but the focus on interpreting “mastery” based on specific items led to the misunderstanding. Dr. Kolstad concludes his paper by describing four options for setting standards that do not rely on RP 67 and that would improve the correspondence between ALDs, test performance, and use of test information. Those options are reproduced in their entirety below.
**Option 1: Use a judgmental process to choose an increment to the IRT b parameter that takes into account the principal policy uses of the data.** For some policy purposes, the balance between false positives and false negatives may differ from those for other purposes. For example, if special services such as literacy remediation are going to be targeted to low performers, we ought to be very sure that they need the services by setting the response probability convention below 0.50 rather than targeting setting it above 0.50 to ensure that examinees are more than capable of the tasks they are set. Since this is a policy decision that depends on the purpose for which the data are expected to be used, there is no reason to rely on the conventional practice of a 0.65 probability. However, it is difficult for assessment programs with multiple uses to focus its procedures on any one expected use. This approach also uses different criteria to place cognitive items and examinees on an assessment scale and produces a degraded correspondence between the percentage of questions answered correctly and the percentage of items that meet the response probability convention.

**Option 2: Map items by matching the distribution of scores in the population and the p-value of the item.** By assigning to the test question the scale score corresponding to the point on the latent distribution at which the percentage of the population achieving at least that point matches the percentage of the population that answers the question correctly. This approach matches the population distributions of success on cognitive items and success at points along the proficiency scale. However, this method uses different criteria to place cognitive items and examinees on the assessment scale, places more stringent standards on easy questions and less stringent standards on harder questions (compressing the items together along the scale). When dealing with nontechnical, content-expert panelists, this method might need an explanation to understand how the difficulty of the individual items corresponds to scale scores.

**Option 3: Map items using simple response patterns.** Under this option, each cognitive question would be assigned the scale score that would be received by an examinee if that question were the most difficult item answered correctly in a simple scalogram pattern of responses. This approach produces a good match between the mapping of items and the percentage of correct answers needed to qualify for the corresponding score. However, this method places less stringent standards on easy questions and more stringent standards on harder questions spreading the items out along the scale. It uses similar, but not identical criteria to place cognitive items and examinees on the assessment scale. In my view, this method has an intuitive explanation that helps nontechnical panelists to understand why the probability of a correct response for an item on the margin is close to 0.50, yet the examinee possesses the ability to answer that or a similarly difficult question correctly.

**Option 4: Map items at the IRT threshold parameter (without adjustment for guessing).** This approach uses the same criteria to place cognitive items and examinees on the scale, places equally stringent standards on all items, produces a passable correspondence between the percentage of questions answered correctly and the percentage of items that meet the response probability convention, but when dealing with nontechnical, content-expert panelists, will need an explanation to counter the basic intuition about the lack of predictability about success with the individual items that they are responsible for examining closely.

**Discussion**

The agenda included seven topics to guide the discussion. Dr. Sunny Becker (HumRRO) facilitated a discussion around these topics and related ideas. However, the conversation appropriately meandered among and across the topics. Rather than forcing this summary to the
original discussion topics, ideas are presented here along the following themes that arose organically:

- Proposed revisions to the Board policy on achievement level setting
- Purposes of reporting ALDs
- Audiences for reporting ALDs
- Various item ordering methods for ALD development
- “Top down” and “bottom up” ALD development
- Panelists and procedures for validating ALDs and developing reporting ALDs
- Special considerations for math and reading reporting ALDs
- Steps to validate alignment of current ALDs and develop reporting ALDs
- Validation/vetting of reporting ALDs

We include discussion within each area and some limited repetition for context and continuity.

**Proposed Revisions to the Board Policy on Achievement Level Setting**

As noted earlier, Dr. Rosenberg shared select excerpts from the proposed revision of the Governing Board policy on developing achievement levels for NAEP. The Experts discussed the language of these revisions. They unanimously endorsed changing the policy labels to *NAEP Basic, NAEP Proficient,* and *NAEP Advanced* to differentiate the NAEP achievement levels from the use of basic, proficient, and advanced in state accountability systems, as well as other assessment programs.

The Experts stated that the proposed wording of Principle 4 – Periodic Review of Achievement Levels was confusing, specifically “past and recent administrations of NAEP assessments”. They suggested substituting “current” or “recent” for “past and recent.” Awkward wording aside, the Experts generally agreed with the stated frequency of revisiting ALDs “[a]t least once every 10 years or 3 administrations of an assessment, whichever comes later.”

The Experts supported the proposal for reporting ALDs to describe what students *do know and can do* and for the ALDs used in achievement level setting to continue to describe what students *should know and be able to do.*

**Purposes of Reporting ALDs**

The Experts agreed that the purpose of producing reporting ALDs is to support accurate, credible, and defensible statements about NAEP findings reported in the *NAEP Basic, NAEP Proficient,* and *NAEP Advanced* achievement levels adopted by the Board. Further, the process to craft these reporting ALDs should inform the evolution of NAEP into the future, maximizing coherence and consistency as frameworks, item pools, and the technology of assessment evolve. This should include defensible standard setting, as well as periodic review of established cut scores.

These purposes imply more specific requirements. Alignment should be integral among frameworks, item pools, ALDs, and cut scores. The process for creation and review of reporting
ALDs must clearly and accurately communicate to multiple stakeholder audiences with overlapping but nonidentical information needs, and should forestall foreseeable misinterpretations, as feasible.

**Audiences for Reporting ALDs**

An important consideration prior to defining the structure, content, and specificity of reporting ALDs is to determine who the intended audiences are. While the group did not reach consensus on this, the Experts did note that multiple audiences may be interested in reporting ALDs, and they engaged in some discussion regarding whether different audiences might require different information. Potential audiences include state assessment directors, policy makers, and governors’ chiefs of staff. Some of the Experts suggested, and others disagreed, that teachers, parents, and students might be considered target audiences.

Once the Governing Board identifies its intended audiences, the list of audiences could inform the best definitional structure and level of specificity for the reporting ALDs. The Experts suggested that a worthwhile step in advance of developing the reporting ALDs would be a study to present various versions of sample reporting ALDs to representatives of the target audiences, and then testing the audiences regarding their understanding of the ALDs. Further, a step following development of the actual reporting ALDs might be to vet them with a sample of target audience members to ensure they are clear and do not result in misinterpretations. The Experts cautioned against creating a complex system with separate reporting ALDs for specific audiences.

**Various Item Mapping Methods for ALD Development**

To validate ALDs through anchoring studies, NAEP has typically used an item mapping method with an RP consistent with the one used in standard setting. In doing so, there is uniformity between the two tasks and in how the panelists think about the item. Currently, the RPs are set around 0.67, although reading and mathematics exemplar item selections in the past used an RP value of 0.50. Huynh (2000a & 2000b) suggests that RP 0.67 represents the maximum amount of item information for correct responses. As each NAEP Technical Report has pointed out, the maximum total information, for both correct and incorrect responses, is represented by an RP of 0.50. Other researchers found RP 0.67 to be a useful criterion for mastery because panelists can interpret it (e.g., NRC, 2006), as the response probability where 2/3 of students with a given cognitive score level would answer a question correctly. The item maps and exemplars available to the public are consistent with the definitions and descriptions of items determined by a high RP value, although the RP typically ranges from 0.65 – 0.74 depending on item type (e.g., see https://www.nationsreportcard.gov/itemmaps/?subj=MAT&grade=4&year=2017).

As described in Dr. Kolstad’s paper and presentation, high RPs have disadvantages. Different RPs yield different item locations and different cut scores. He demonstrated that mapping by a high RP value will lead to decreased false positive results (for students who may not have the skill, but still get credit for having knowledge of it) but at the same time will lead to increased false negatives (for students do have the skill, but do not get credit for having that knowledge). Moreover, he provides examples of how lay audiences misinterpret what the RP level of an item means to the underlying scale, item difficulty, total score, and the item. He believes that panelists can interpret RP 0.50 as the response probability of the most difficult item in a score-equivalent pattern of responses in which all items below it are answered correctly and all above it are answered incorrectly.
The group noted that regardless of what method and criteria are used to map items to the score scale, there are likely to be some items that do not map into an achievement level because they do not adequately differentiate students at one level from those at adjacent levels. The Board will need to decide how to handle misaligned items both in terms of the development of reporting ALDs and in terms of future administrations of the assessment. Some Experts suggested removing items from the current pool that do not match the ALDs or do not map to the cut score range for which the content matches the ALD. A large percentage of misaligned items could indicate a larger validity issue that would warrant additional study. The Experts noted that while items that map in the below Basic range would not be used in ALDs, they are essential to measuring a broad range of knowledge and skills.

Developing reporting ALDs should be focused on accurately reflecting what students in each category can do. Knowledge, skills, and abilities (KSAs) may be summarized for students right at the cut score, in the middle of the achievement level range, or at the very top of the achievement level range. In addition, the Experts discussed how reporting ALDs would be most useful if the descriptors included specific information on what most students at a particular level can do, what many students can do, and what some students can do in an achievement level range (e.g., Basic, Proficient). If NAEP results include reporting ALDs, score interpretation guides need to be developed to support appropriate score use.

This prompted the Experts to discuss a variety of methods and their rules for item mapping, since item location and ultimately, the ALD interpretation, will be impacted by the chosen method. The approach could change, based on the use and/or users of the reporting ALDs.

The Experts focused their discussion on these approaches:

- Map items at the IRT threshold parameter (without adjustment for guessing), such as RP 67. This is the current practice. The standard setting panelists are asked to think about where a barely qualified examinee has a 2/3 chance of getting the item correct. This is also described as the response probability of where 2/3 of minimally competent students would answer a question correctly. Once the standards have been set, the items can be sorted in the performance levels based on their parameters. The Experts noted that, once the content and skills of the items are described, the fact that the original order was based on an RP value is often lost.

- Map items using simple response patterns (Dr. Kolstad’s preferred method) using a scalogram approach. This method relies on Guttman scaling or cumulative scaling that is often used to measure attitudes. Unlike item response theory, it is not probabilistic. The scale is a unidimensional continuum to help stakeholders predict item responses knowing the total (cumulative) score. However, scales are rarely perfectly cumulative, thus requiring scalogram analyses. In education, the scales are often used to obtain formative information against an expected learning progression. Guttman charts are frequently used in formative assessment because they portray what students know and can do. Guttman scaling has been used in NAEP research to support reporting domains (Schulz & Lee, 2002). To develop the ALDs, patterns of right and wrong responses could be compared. The interpretation becomes easier because it is consistent with stakeholder’s instinctive understanding of higher and lower scores.

- Establish conditional p-values based on the students within any achievement level, such as Proficient. This method removes the items from their underlying scale. The p-values could be conditioned at 80% proficient, for example, to describe what most students in this achievement level can do, or at 60%, to reflect what many of the students in the
achievement level can do. These p-values can be used to identify items that exemplify what students within that category know and can do with different levels of certainty. This would result in a good description of the items that students within a group can do with different levels of probability/certainty, and this can be the basis for the ALDs.

The pros and cons of using RP values for mapping were discussed. The Experts did not come to consensus on whether to continue to approach the creation of ALDs based on a specific RP value (e.g. RP 67). There was concern that electing not to base item sets on RP 67, when that RP value was used during standard setting might create inconsistency in the overall system. Some of the Experts also endorsed the idea that “mastery” meant that students necessarily had more than a 50% chance of answering an item correctly. Others thought “mastery” was a flawed interpretation of a high RP value. Some of the Experts concluded that validation of ALDs based on items selected from an RP value designated for mastery is challenging and may lead to spurious conclusions based on skewed misclassifications.

Since the Governing Board’s goal for reporting ALDs is to describe what students within a group know and can do, it may be more straightforward to eliminate RPs from consideration. Therefore, some Experts suggested using conditional p-values. The idea of mastery based on item content led to a discussion about the definition of mastery that should be applied to NAEP. There was some agreement that a small-scale study should be considered to attempt to create ALDs based on conditional p-values of items among students within each NAEP classification. However, not all Experts endorsed this idea.

“Top Down” and “Bottom Up” ALD Development

The Experts noted that the ALDs which are used by standard setting panels and item developers are developed in a top-down way. That is, policy definitions express what each achievement level means in a general, “high-level” sense, without reference to specific content areas or grade levels. These general guidelines drive interpretation of a framework for a specific content area and grade level to describe in greater detail what a student should be able to do. For example, Figure 4 shows the proposed revised policy ALD for NAEP Proficient performance and examples of ALDs that could be used to set achievement levels for the NAEP Proficient level at grades 4 and 8. Currently, these ALDs serve as the reporting ALDs for NAEP mathematics.

On the other hand, developing reporting ALDs based on the items that empirically map onto each achievement level (Figure 5) was described by the Experts as a “bottom up” process. The item information provides a smaller grain size than the frameworks, so summaries of what students actually know and are able to do within each achievement level are likely to be at a greater level of detail than summaries developed from the framework objectives. Any given item pool is necessarily a limited subset of the possible items posited by the framework, so the “bottom up” process would inevitably contain some “holes.” Developing reporting ALDs in this way, and then comparing those to the ALDs depicted in Figure 4, offers a rigorous confirmation that the frameworks, items, and cut scores, are all highly aligned. If the reporting ALDs contradict the ALDs used to set achievement levels, or if some of the specific statements are missing from one or the other, reconciliation would be necessary. Any bottom-up description cannot cover all that the framework intends. Small adjustments would not be problematic, but substantial discrepancies would require a deeper investigation and policy decisions.
Proposed Policy ALD: NAEP Proficient

This level represents solid academic performance for each NAEP assessment. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real world situations, and analytical skills appropriate to the subject matter.

Content ALD: Grade 4 Mathematics Proficient:

Fourth-grade students performing at the Proficient level should consistently apply integrated procedural knowledge and conceptual understanding to problem solving in the five NAEP content areas.

Fourth-graders performing at the Proficient level should be able to use whole numbers to estimate, compute, and determine whether results are reasonable. They should have a conceptual understanding of fractions and decimals, be able to solve real-world problems in all NAEP content areas, and use four-function calculators, rulers, and geometric shapes appropriately. Students performing at the Proficient level should employ problem-solving strategies such as identifying and using appropriate information. Their written solutions should be organized and presented both with supporting information and explanations of how they were achieved.

Content ALD: Grade 8 Mathematics Proficient:

Eighth-grade students performing at the Proficient level should apply mathematical concepts and procedures consistently to complex problems in the five NAEP content areas.

Eighth-graders performing at the Proficient level should be able to conjecture, defend their ideas, and give supporting examples. They should understand the connections between fractions, percents, decimals, and other mathematical topics such as algebra and functions. Students at this level are expected to have a thorough understanding of Basic level arithmetic operations—an understanding sufficient for problem solving in practical situations. Quantity and spatial relationships in problem solving and reasoning should be familiar to them, and they should be able to convey underlying reasoning skills beyond the level of arithmetic. They should be able to compare and contrast mathematical ideas and generate their own examples. These students should make inferences from data and graphs, apply properties of informal geometry, and accurately use the tools of technology. Students at this level should understand the process of gathering and organizing data and be able to calculate, evaluate, and communicate results within the domain of statistics and probability.

Figure 4. Top-down development of content ALDs from higher level policy ALDs.
Content ALD: Grade 4 Mathematics Proficient:
Fourth-grade students performing at the Proficient level can consistently apply integrated procedural knowledge and conceptual understanding to problem solving in the five NAEP content areas.

Most fourth-graders performing at the Proficient level can use whole numbers to estimate, compute, and determine whether results are reasonable. Many have a conceptual understanding of fractions and decimals; be able to solve real-world problems in all NAEP content areas; and use four-function calculators, rulers, and geometric shapes appropriately.

Content ALD: Grade 8 Mathematics Proficient:
Eighth-grade students performing at the Proficient level can apply mathematical concepts and procedures consistently to complex problems in the five NAEP content areas.

Most eighth-graders performing at the Proficient level can conjecture, defend their ideas, and give supporting examples. Many can understand the connections between fractions, percents, decimals, and other mathematical topics such as algebra and functions. Some students at this level have a thorough understanding of Basic level arithmetic operations—an understanding sufficient for problem solving in practical situations.

Figure 5. Bottom-up development of hypothetical reporting ALDs from items that map onto a given NAEP achievement level.
Panelists and Procedures for Validating ALDs and Developing Reporting ALDs

The discussion of who should serve on panels to validate existing ALDs and create reporting ALDs began with a review of the panelist qualifications for NAEP achievement level setting activities. The proposed revised policy states that achievement level setting panels shall consist of at least 50% teachers, with non-teacher educators (e.g., curriculum directors, academic coaches, principals) accounting for no more than half the number of teachers. The remaining panelists would be non-educators who represent the perspectives of additional stakeholders representing the general public, including parents, researchers, and employers. Panelists should have expertise and experience in the specific content area in which the levels are being developed; expertise and experience in the education of students at the grade under consideration; a general knowledge of assessment, curriculum, and student performance; and shall reflect diversity in terms of gender, race/ethnicity, region of the country, urbanicity, and experience with students with disabilities and English language learners.

The Experts agreed that the achievement level setting panelists should not also create reporting ALDs, since the tasks are different, both activities are very time-intensive, and it may be confusing to perform multiple tasks with different instructions and context. There was less consensus on who should serve on panels to validate existing ALDs and develop reporting ALDs. There are benefits to including participants from the framework committee; they come to the task prepared with knowledge of the framework and rationale behind its development. They discussed mixed groups including some members of the framework committee and some new individuals who would have been qualified to serve on the framework committees but did not serve. The Experts felt that it would not be appropriate to include individuals who lacked content expertise in the subject.

In terms of the number of panelists, the Experts suggested including up to five or six panelists per grade and subject. With six panelists, two groups of three can divide the work or replicate the work in each grade/subject. For math, this would mean that there would be approximately 15-18 panelists across grades 4, 8, and 12. The panelists could sit at tables of 3 for some activities and then be grouped as tables of 6 for other activities.

In terms of procedures, the Experts proposed that panelists would need to engage in a validation step to compare results from an item mapping (or similar) procedure to the ALDs currently in use. Most Experts endorsed a strong design where panelists would write new ALDs based on the item mapping procedure. Then the new ALDs based on the empirical data would be compared to the ALDs currently in use (and in most cases, the ALDs that were used to set the achievement levels). Consistency checks between the two sets of ALDs would identify serious mismatches and serve as a validity check; this would effectively be a comparison of the ALDs produced from a "top-down approach" with ALDs produced from a "bottom up" approach. Training and facilitation are very important. The Experts noted that it is challenging to distill ALDs from item sets, and that panelists must be adept at recognizing when items may perform unexpectedly due to idiosyncratic reasons. The panelists would require training on how to take individual items and develop ALD statements, and specific examples would be needed.

Some Experts argued that it might not be necessary to write new ALDs to perform this validation step; instead panelists could be asked to judge the extent to which the current ALDs could be verified using the item mapping data. Other Experts were concerned that this procedure might be more open to bias and that it was important for NAEP to use the strongest design possible to maintain its reputation for being the gold standard, especially when conducting the studies for math and reading. The wording of the first recommendation in the evaluation implies that the
designs used for previous NAEP anchoring studies in reading and math are acceptable, and those studies did use the stronger design of writing new ALDs to compare to the existing ALDs.

Another option that was discussed but largely dismissed was asking panelists to look at items and predict which achievement level they should represent, and then compare those results to the actual results from an item mapping procedure. Several Experts noted that based on their experiences, it is very difficult for panelists to estimate the ALD targets for items.

Following the validation step, the Experts agreed that panelists would need to use the item mapping data to draft statements for reporting ALDs. For example, if the reporting ALDs are written in terms of the things that most/many/some students in each achievement level can do, then panelists would need to examine items and write descriptions based on the percentage of students who correctly answered (or would be expected to correctly answer) various items. Some Experts recommended using statements about what most/many/some students can do rather than what the “typical” student can do. The use of items for creating reporting ALDs should be limited to recent administrations of the assessment. Reporting ALDs should reference specific content and should provide users with as much specificity as they require, within the limits of good measurement practice.

If all of these steps were included, the meetings to validate ALDs and create reporting ALDs would be likely to take at least three or four days.

**Special Considerations for Math and Reading Reporting ALDs**

Dr. Rosenberg began a discussion of special considerations for developing reading and mathematics ALDs given the first recommendation of the recent evaluation of NAEP achievement levels. She noted because NAEP has transitioned to digital administration, these data currently exist for grades 4 and 8 in 2017; however, the grade 12 reading and mathematics assessments will not be administered digitally until 2019, so those data will not be available until late next year. She explained that NAEP’s transition to a digitally-based administration involves “trans-adapting” the items from the paper-pencil to a digital assessment version, resulting in item parameters not being on the same scale. There was consensus among the participants there must be coherence among the reporting ALDs across the three grades; they believed there would be too much risk for inconsistency if the reporting ALDs were developed at different times. The participants suggested developing reporting ALDs at the same time for the three grades but using data from the grade 12 paper-pencil administration for a pilot study to develop draft reporting ALDs and then creating operational reporting ALDs when the grade 12 digital data become available.

Dr. Rosenberg raised concerns about being able to complete all activities so the newly developed reporting ALDs could be used to report the 2019 NAEP results. The Experts noted the tendency for the Governing Board to be deliberate when making changes to the NAEP program. They felt it was not appropriate for activities to be rushed to address the recommendation from the evaluation. The Experts felt strongly the Governing Board should conduct a feasibility study that includes one grade for each content area and a pilot study of the new reporting ALDs at all three grades prior to using them operationally. Additionally, they recommended that the Governing Board plan strategically so there will be sufficient time between the pilot study and operational study to make appropriate changes. The Experts suggested that the Governing Board conduct a special study to examine the efficacy of the new reporting ALDs and to determine their usefulness to the various stakeholders. They also suggested that the Governing Board conduct focus groups to obtain feedback about the use and interpretation of “most/many/some” phrasing in the new reporting ALDs. The Experts
suggested that the Governing Board staff prepare a recommended timeline to develop the new reporting ALDs, along with a rationale that includes the validity evidence that will be collected from each activity, so that Board members can appreciate what is needed to produce reporting ALDs that are useful and widely understood.

**Steps to Validate Alignment of Current ALDs and Develop Reporting ALDs**

Dr. Becker led the group through development of an explicit list of suggested steps to verify alignment between the item pools, cut scores, and ALDs and to develop new reporting ALDs for mathematics and reading at grades 4, 8, and 12:

1. Governing Board defines users of the NAEP reporting ALDs.
2. Conduct focus groups with users of the NAEP reporting ALDs to determine what they find useful.
3. Issue a Request for Proposal (RFP).
4. Prepare a design document for validating the alignment of items, cut scores, and ALDs and for developing new reporting ALDs; the design will include pilot grade 12 reporting ALDs based on paper-pencil data and operational ALDs based on the 2019 digitally-based administration.
5. Conduct a feasibility study that includes one grade each at reading and mathematics.
6. Conduct a pilot study at each of the three grades for reading and mathematics.
7. Evaluate the results from the pilot study, including by sharing the resulting reporting ALDs with potential users; if major changes are made to the process, then conduct another pilot study.
8. Use this process operationally in 2021 (using 2019 data rather than waiting for 2021 data) to report NAEP results for reading and mathematics.
9. Develop a process to evaluate and vet the reporting ALDs.
10. Develop communications and dissemination plans using the new reporting ALDs.

The Experts discussed the importance of clarity around the different types and sequencing of ALDs, including when ALDs and cut scores are adopted by the Governing Board. At this point, an ALD alignment study must include decisions and actions regarding any items determined to be misaligned. The subsequent development process for reporting ALDs should include an anchor study/item mapping and an evaluation of alignment between reporting ALDs and content ALDs. Adherence to this formal process would culminate in reporting ALDs representing a coherent and consistent system of frameworks, item pools, ALDs, and cut scores.

**Validation/Vetting of Reporting ALDs**

The Experts discussed validation of reporting ALDs in terms of their accuracy and utility. They noted that NAEP is often considered to be the “gold standard” of assessments and in some ways, provides a methodological approach that may inform state and other assessments. They noted that the process for producing, monitoring, and managing reporting ALDs affords another opportunity to demonstrate best practices to benefit other testing programs.
First, using common item sets should lead replicate panels to come to essentially the same content for reporting ALDs in each of the categories. The Experts also discussed using both RP values and conditional p-values to create reporting ALDs, and then comparing them. However, there was not agreement on the RP value that would be most appropriate (50 or 67), so this method of validation may be more complex.

The Experts also suggested conducting focus groups or “market research” to determine whether the reporting ALDs were understood by users and met their needs. They identified several potential users (e.g. state education agency officials, governor’s office staff, district-level education staff) who might use the reporting ALDs. The vetting of the ALDs would include exploration of potential misuses or misinterpretations of data, as well as gathering feedback on the utility and ease of interpretation of the reporting ALDs.

**Summary and Reflections**

Dr. Rosenberg expressed her appreciation for the Experts’ insights. She highlighted the key take-away points she planned to share with COSDAM at their August meeting. The Experts agreed with this summary.

- **Wording of proposed policy changes**
  - Endorse new labels of NAEP Basic, NAEP Proficient, NAEP Advanced
  - “Past and recent administrations of NAEP assessments” is not clear.
  - Agree with reviewing reporting ALDs every 3 administrations or 10 years, whichever comes later.

- **Articulate the current procedures for verifying that NAEP items are aligned to their frameworks. States are required to have independent reviews to evaluate the alignment between their item pools and frameworks.**

- **What are the interpretation and use arguments for NAEP ALDs?**
  - What sources of evidence are needed?

- **Suggestion to phrase reporting ALDs in terms of what most/many/some can do based on actual student performance.**

- **Use separate panels to set standards and develop reporting ALDs to avoid cognitive shift from one purpose to another.**
  - For setting reporting ALDs, use a mix of panelists from framework committees and educators who were qualified to serve on the framework committee but did not.
  - This process may need 3-4 days.

- **Be planful. Release of reporting ALDs in 2021 is more feasible than in 2019.**
  - Use 2015 grade 12 paper-pencil results and 2017/2019 grades 4 and 8 DBA results. Update grade 12 when 2019 results are available.
  - Pilot test could use 2015/2017 data or 2017 grades 4 and 8 data, depending on how many items are needed.

- **Identify validation steps needed for each part of the process (see list of 10 steps noted above to validate alignment of current ALDs and produce reporting ALDs).**
References


Appendix A: Meeting Agenda, Attendees, and List of Read-Ahead Materials
DAY 1

9:00 – 9:15  Welcome, Introductions, and Meeting Goals  
Dr. Sunny Becker

9:15 – 10:00  Setting the Context  
Dr. Sharyn Rosenberg

10:00 – 10:30  Considerations for Reporting ALDs for NAEP  
Dr. Hillary Michaels*

10:30 – 10:45  Break

10:45 – 11:15  State Approaches to PLD Review and Revision  
Dr. Marianne Perie*

11:15 – 11:45  History of Anchor Studies for NAEP  
Dr. Susan Loomis*

11:45 – 12:15  Methodological Considerations  
Dr. Andy Kolstad*

12:15 – 1:00  Break for lunch

1:00 – 1:15  Review and Revise Discussion Topics

1:15 – 5:00  Group Discussion  (break from approximately 3:00 – 3:15)
- Use and Range of Reporting ALDs
- Methodology
- Panelists and Procedures
- Special Considerations for Math and Reading ALDs

6:00  Meet for optional group dinner

DAY 2

9:00 – 9:15  Review of Previous Day and Plan for Today  
Dr. Sunny Becker

9:15 – 12:15  Group Discussion  (break from approximately 10:30-10:45)
- Special Considerations for Reporting ALDs When Setting New Achievement Levels
- Validation/Vetting of Reporting ALDs
- Recommendations for Special Studies

12:15 – 1:00  Break for lunch

1:15 – 2:45  Group Discussion
- Decision Points for the Governing Board
- Summary of Recommendations and Next Steps

2:45 – 3:00  Wrap-up  
Dr. Sunny Becker

* Session will consist of a brief presentation by an author, reminding the Experts about content of a read-ahead document. This will be followed by clarifying questions.
Attendees

Expert Panelists:
Dr. Susan Davis-Becker, ACS Ventures, LLC
Dr. Karla Egan, EdMetric, LLC
Dr. Ed Haertel, Stanford University
Dr. Steve Ferrara, Measured Progress
Dr. Andy Kolstad, P20 Strategies LLC
Dr. Susan Loomis, Consultant
Dr. Barbara Plake, University of Nebraska-Lincoln
Dr. Laurie Wise, HumRRO

Governing Board Staff:
Ms. Michelle Blair
Dr. Sharyn Rosenberg
Dr. Lisa Stooksberry

HumRRO:
Dr. Sunny Becker
Dr. Monica Gribben
Dr. Hillary Michaels
Dr. Sheila Schultz
Dr. Arthur Thacker

NCES:
Dr. Enis Dogan

ETS (NAEP Design, Analysis, and Reporting Contractor):
Dr. Mary Pitoniak

On the phone (for part of the meeting):
Dr. Marianne Perie, University of Kansas


Read-ahead Materials


A free PDF can be downloaded at: https://www.nap.edu/catalog/23409/evaluation-of-the-achievement-levels-for-mathematics-and-reading-on-the-national-assessment-of-educational-progress


Uses of NAEP  
For August 3, 2018 COSDAM Discussion

Over the past couple of years, COSDAM has had several discussions about the need to explicitly state how NAEP results (in general, and achievement levels in particular) are intended to be used, and then to focus dissemination efforts on increasing the most appropriate and impactful uses of NAEP. The very first standard (Standard 1.0) of the *Standards for Educational and Psychological Testing* states: “Clear articulation of each intended test score interpretation for a specified use should be set forth, and appropriate validity evidence in support of each intended interpretation should be provided” (AERA, APA, & NCME, 2014; p. 23).

The Governing Board’s Strategic Vision includes a goal to expand the availability, utility, and use of NAEP resources, in part by creating new resources to inform education policy and practice (SV #3). COSDAM activities to address this goal include: conducting research on how NAEP results are currently used (both appropriately and inappropriately) by various stakeholders; developing a statement of the intended and unintended uses of NAEP data (in conjunction with NCES); and working with NCES to produce documentation of validity evidence in support of the appropriate uses of NAEP.

One of the major recommendations (Recommendation #5) from the recent evaluation of NAEP achievement levels is: “Research is needed to articulate the intended interpretations and uses of the achievement levels and to collect validity evidence to support these interpretations and uses. In addition, research is needed to identify the actual interpretations and uses commonly made by NAEP’s various audiences and evaluate the validity of each of them. This information should be communicated to users with clear guidance on substantiated and unsubstantiated interpretations” (National Academies of Sciences, Engineering, and Medicine, 2017, p. 13). The proposed revised policy on Developing Student Achievement Levels for NAEP references an interpretative guide that would accompany NAEP reports and include specific examples of appropriate and inappropriate interpretations and uses of the results (Principle 3h).

As part of the Technical Support contract, the Human Resources Research Organization (HumRRO) has been conducting research on how NAEP results have been used by various audiences, including: federal, state, and local policymakers; educators; media; education researchers; and the general public. The first phase of this work (currently underway) is to analyze existing artifacts produced by these various audiences; a potential follow-up activity is to
conduct interviews and/or focus groups to gather additional information that cannot be gleaned from existing artifacts, if warranted.

As part of their work conducting the evaluation, the National Academies of Sciences, Engineering, and Medicine (2017) also conducted some research on how the achievement levels are being used; the evaluation report includes a summary of uses, interpretations, and actions for the NAEP achievement levels (p. 192-193).

Using preliminary findings from the research efforts referenced above, along with their own knowledge of common uses and interpretations, Governing Board staff developed two high level lists to support the August COSDAM discussion. The first is a list of primary uses (how different types of NAEP results are used); the second indicates secondary uses (common interpretations and actions based on those uses). The lists do not attempt to differentiate appropriate versus inappropriate uses and interpretations.

During the upcoming August Board meeting, COSDAM members will discuss how to use this information to inform next steps for: 1) developing a statement about appropriate uses of NAEP; and 2) developing an interpretative guide for communicating how the NAEP achievement levels should be used.

**Discussion Questions**

- What are the general principles and considerations for developing a statement of appropriate and inappropriate uses of NAEP?
- What are the general principles and considerations for developing an interpretative guide for communicating achievement level results?
- In order to develop these documents, is it necessary to gather additional information about how NAEP is used by conducting interviews and/or focus groups? If so, what are the priority audiences and questions to be answered?

**References**


Common Uses, Interpretations, and Actions Based on NAEP Data

Primary Uses

- Compare NAEP scale scores and/or achievement levels at a **single point in time** across states, districts (TUDA), and/or student groups
- Compare NAEP scale scores and/or achievement levels **over time (trends)** for the nation, states, districts (TUDA), and/or student groups
- Rank order states or districts in terms of NAEP scale scores and/or achievement levels overall and/or for a specific student group
- Analyze **performance gaps** in NAEP scale scores and/or achievement levels between two student groups at a single point in time
- Analyze **changes in performance gaps** in NAEP scale scores and/or achievement levels between two student groups over time (gap trends)
- Validate performance or changes in performance on state tests
- Analyze the relationship between contextual variables and NAEP scale scores and/or achievement levels
- Describe the context in which students learn from information gathered by student, teacher, and school questionnaires
- Compare NAEP scale scores and/or achievement levels **across subject areas**
- Compare NAEP scale scores and/or achievement levels **across grades**
- Compare NAEP scale scores and/or achievement levels **before and after a program or policy is implemented**
- Estimate the percentage of students who are **academically prepared for college** by the end of high school
- Show examples of what students know and can do through sample items and item maps
- Establish a **common scale for linking state tests** and comparing results across all school districts (e.g., Stanford Education Data Archive)
- Link other assessments to NAEP to provide state-level results on other assessments that were not administered at the state level (e.g., TIMSS)
- Establish a common scale for comparing the rigor of state standards to each other and to NAEP Proficient
- Compare the percentage of students at or above each achievement level on NAEP and on other assessments, including state and international assessments
- Serve as a benchmark of performance at NAEP Proficient to inform standard settings on other assessments
Secondary Uses

• To evaluate whether current programs and policies are effective
• To support the need for new programs and policies
• To influence decisions about funding for educational policies and programs
• To influence legislation
• To determine whether the nation, states, and/or TUDAs are making progress for students overall and/or selected student groups
• To evaluate the quality of education at a single point in time and/or over time
• To claim that some states and/or districts are doing a better job educating students based on their rankings on NAEP
• To identify where there are large performance gaps and/or interventions are needed
• To identify states and/or TUDAs who are doing something extraordinary so that best practices can be shared
• To criticize states for lying about the percentage of students at or above Proficient if it varies substantially from NAEP
• To generate and test hypotheses about factors related to student achievement (education research)
• To claim that students should do more of X because X is correlated with higher performance
• To determine whether U.S. students will be internationally competitive
• To call for higher standards
• To call for more accountability systems
• To claim that the majority of students lack basic skills (or are faring well)
• To make claims about the percentage of students who are performing “on grade level”
• To inform the development of state content standards
Memorandum #1: Considerations Related to the Validation of NAEP Achievement Levels

Prepared for:
Sharyn Rosenberg
National Assessment Governing Board
800 North Capitol Street N.W., Suite 825
Washington DC 20002
Phone: (202) 357 6940
Email: Sharyn.Rosenberg@ed.gov

Prepared under:
ED NAGB 17 C 0002
Contracting Officer: Munira Mwalimu
Phone: (202) 357 6906
Email: Munira.Mwalimu@ed.gov

Authors:
Arthur A. Thacker
Tanya Longabach

Date:
February 5, 2018
Memorandum #1: Considerations Related to the Validation of NAEP Achievement Levels

Table of Contents

Introduction ................................................................................................................................ 1
Summary of Achievement Levels Use and Interpretation .......................................................... 2
Inferences from Various Stakeholders ....................................................................................... 5
  Policymakers ...................................................................................................................... 5
  Educators ........................................................................................................................... 5
  Researchers and Business Leaders ................................................................................... 7
  The Media .......................................................................................................................... 9
  The General Public ............................................................................................................10
Approaching Validation of the NAEP Performance Levels Using a Validity Argument ..........11
  Contextual Factors that Represent Challenges for Constructing a Validity Argument for NAEP ALDs ........................................................... 15
  Using NAEP Achievement Levels to Inform Statewide Testing Standards ......................16
  Evaluating NAEP’s Achievement Levels for an Evolving Educational Landscape ..............17
  Summary: Steps Toward Developing and Validity Framework for NAEP Achievement Levels .................................................................................................................................17
References ...............................................................................................................................19

List of Tables

Table 1. Highest and Lowest Performing States on NAEP Reading and Mathematics Grades 4 and 8 ................................................................................................................... 6
Table 2. Test Design Claims, Assumptions, and Evidence .......................................................14

List of Figures

Figure 1. Relationships among theory of action (TOA), interpretive argument, and validity argument .................................................................................................................................11
Memorandum #1: Considerations Related to the Validation of NAEP Achievement Levels

Introduction

One common characteristic of educational assessments is the need to make broader inferences about students' knowledge and abilities from specific behaviors (Mislevy, Almond, & Lukas, 2003). Since we cannot directly see the knowledge and abilities we wish to measure, or to observe them in full, our measurement of those constructs is a proxy measurement. Therefore, we need to justify the inference that the observable behavior is a manifestation of the unobservable construct we are trying to measure. The ways that we interpret the score that students receive on an assessment depends on the inferences we make between the observed student behavior and the unobserved construct.

Validity is a property of the interpretations assigned to scores, and these interpretations are considered valid if they are supported by convincing evidence. In order to evaluate the plausibility of a test score interpretation, it is necessary to be clear about what the interpretation claims. That is, a claim should be made explicitly and directly about the inferences we intend to make. The interpretive argument specifies a network of inferences leading from the scores to the conclusions we intend to make based on those scores, as well as the assumptions supporting these inferences. In assembling and organizing evidence for the interpretive argument, we are developing a validity argument, the goal of which is to show that the interpretive argument is plausible (Kane, 2001). The process of developing the validity argument is known as validation. If the proposed interpretation of test scores is limited, as it is for some observable attributes, the requirements for validation can be very modest. If the proposed interpretations are more ambitious, as they are for traits and theoretical constructs, more evidence and more kinds of evidence are required for validation (Kane, 2013).

The Standards for Educational and Psychological Testing (AERA et al., 2014) place great importance on validity, calling it "the most fundamental consideration in developing tests and evaluating tests" (p.11). Specifically, Standard 1.0 states that "clear articulation of each intended test score interpretation for a specific use should be set forth, and appropriate validity evidence in support of each intended interpretation should be provided" (p.23). The associated standard cluster 1, including standards 1.1-1.7, elaborate on various aspects of validity that are essential to support assessment uses and interpretations.

Argument-based validation, as described by Kane (2006; 2013), primarily involves supporting the intended inferences that can be drawn from assessment scores. We typically begin by identifying the persons or groups that are expected to draw inferences from the test scores and we then describe those inferences in as much detail as possible. Once we understand the expected inferences, we can generate evidence to support the use of the test scores for those specific purposes. The National Assessment of Educational Progress (NAEP) is a very complex assessment system that does not produce individual students’ scores. Many of the inferences that NAEP supports are quite different from most other student assessments.
The National Assessment Governing Board’s (Governing Board) recent Strategic Vision identifies policymakers, educators, researchers and business leaders, the media, and the general public as stakeholders who are expected to use NAEP results. The Strategic Vision is not so specific as to describe how each group is expected to use NAEP results, but it does indicate that they should be informed “about what America’s students know and can do in various subject areas and compare achievement data over time and among student demographic groups.” The Strategic Vision also states that NAEP should “inform education policy and practice.”

The Governing Board is working towards developing a statement of intended and appropriate uses for both scale scores and achievement levels. HumRRO is currently conducting a research study to determine how various audiences have used and interpreted NAEP results. However, the current lack of specificity in the inferences each indicated group might make represents a substantial challenge for validation. For that reason, we will approach the creation of this section of the validity argument in two ways. First, we will address some of the most straightforward interpretations of NAEP results. These interpretations are well-described on the website and are most commonly associated with the Nation’s Report Card. We will not provide an exhaustive list of these interpretations and inferences here, but we will demonstrate a claim structure that might be used to support them. Then we will seek out inferences the identified groups have actually made from NAEP results. We will then describe how those inferences were supported and discuss additional claims and evidence that might be necessary for validation of those inferences.

Note that this memorandum is not comprehensive. Our goal is to provide guidance on how NAEP achievement levels might be validated for making specific inferences. The number of potential inferences that might be made and the amount of documentation available to potentially support those inferences is well beyond the scope of this memorandum. The examples we include in this memorandum, while important, do not necessarily represent the most important validation issues or interpretations of NAEP levels rather, they were chosen to be illustrative of the range of inferences. Where possible, we summarize the literature related to common claims, but these summaries do not represent an exhaustive literature review.

Summary of Achievement Level Descriptors Use and Interpretation.

Achievement level descriptors (ALDs) are the descriptions of knowledge, skills, and abilities of students at specific achievement levels. ALDs often include input from policymakers, stakeholders, and content experts. Egan, Schneider, and Ferrara (2012) identify three major uses of ALDs: standard setting guidance, test development, and score interpretation.

Some researchers identify standard setting as a primary use of ALDs. For example, Bourque (2000) said that the most important function of ALDs is considered to be providing “a mental framework or structure for standard setting panelists” (p.8). The clarity of ALDs is essential for setting meaningful cut scores (Kane, 2001): if ALDs are unclear, panelists cannot confidently determine how to sort examinees into groups based on achievement and set the cut scores. ALDs highlight what examinees need to accomplish to meet performance standards (Hambleton, Pitoniak & Copella, 2012).

---

Using ALDs to guide test development has been a topic of some debate. Some researchers suggest that ALDs can be used as a tool to guide the development of test blueprints, item specifications, and items themselves (Egan, Schneider, & Ferrara, 2012). While this idea makes sense, it is predicated on the ability of item writers to not only make judgments regarding the specific content that the item assesses, but also of the item difficulty, so that a wide range of items can be created that probe different ability levels as described in the ALDs. This use of ALDs may be challenging until it becomes clearer what factors affect item difficulty (Schneider, Huff, Egan, Tully, & Ferrara, 2010).

ALDs are an essential instrument of score interpretation; they were introduced in NAEP standard setting with the specific goal of making scale score interpretation easier and more meaningful (Kane, 2001; Bourque, 2009; Egan, Schneider, & Ferrara, 2012). Referencing performance categories (e.g. advanced, proficient, basic) used to divide a score reporting scale into ordered score intervals – rather than referencing the test scores themselves – may be a more understandable way of communicating test results (Hambleton, Pitoniak & Copella, 2012). With ALDs providing the descriptions of what the students at each of the performance categories know and can do, the stakeholders can easily see what abilities are associated with a scale score. ALDs give meaning to the cut scores established during a standard setting session.

National Academies of Sciences, Engineering, and Medicine (2017) outline the following purposes for having achievement standards:

- to be able to summarize students’ present achievement and track their progress;
- to mark disparities between what we expect students to know and what they actually know;
- to stimulate policy conversations about educational achievement (and possibly discussions about methods of achieving the levels we want the students to be at);
- to identify content areas of high and low performance, as well as student subgroups of high and low performance; and
- to inform policy interventions and reform measures to improve student learning.

These uses of ALDs can at times be challenging to reconcile (Egan, Schneider, & Ferrara, 2012). For example, when ALDs are first created prior to a standard setting (so they can guide standard setters), they may be mainly aspirational; that is, they may articulate the policymaker’s vision of the goals and rigor of achievement and answer the question “what should the students at specific achievement levels know and be able to do?” Later on, after the assessment data are collected and student scores are being reported by proficiency levels, the question being answered may change to “what do the students actually know?”

The validity of the assessment score inferences and ALD validity are interrelated. In an ideal situation, ALDs would guide the development of the test, so that the test is aligned with the construct of interest. The ALDs describe the degree to which students at each performance level possess this construct. The ALDs could then guide item writers in creating items that are aligned with this construct and elicit the knowledge that is aligned with the construct of interest. ALDs could also guide standard setters so they create cut scores with the same construct concept in mind as the item writers. Because the test is aligned with ALDs, and ALDs describe the degree to which the student possesses the construct of interest, the test assesses appropriate content. The ALDs used in score reporting, in turn, are aligned with test items and
represent the observed skills of students at a particular performance level. However, this process is seldom followed in reality (Egan, Schneider, & Ferrara, 2012). The disconnects between ALDs, cut scores, and the assessment itself, including assessment framework, items, and scoring, at different stages of the process may challenge the validity of ALDs.

Answers to the following questions would support the validity of the standards.

- Are the standards reasonable (based on a common understanding of what students should know and be able to do in the subject area)?
- Are the standards informative to the public?
- Can the public understand what students are expected to know and do?
- Do the standards lead to appropriate interpretations?

These general and typical purposes described above are consistent with the intended purposes of the NAEP ALDs as described in the Governing Board’s Strategic Vision. The typical questions asked as part of the validation of standards are also applicable to the NAEP ALDs. After reviewing information related to the creation and use of the NAEP ALDs, we identified several issues that may represent challenges for their validation. These include:

- There is disagreement and/or confusion among stakeholders about how to interpret the meaning of “proficient” described by the NAEP ALDs.
- There has been disagreement from the beginning of NAEP administration regarding what the achievement levels should be; they have been declared “trial” and continue to have this status.
- The achievement levels are considered to be unreasonably high by some people.
- There is little guidance on how the achievement levels should be used and interpreted.

Our summary is very similar to validation challenges described by National Academies of Sciences, Engineering, and Medicine (2017): It remains challenging to find guidance on the intended interpretations and uses of NAEP achievement levels for stakeholders, including educators, administrators, and the public. The support for the uses of the achievement levels—the way that NAEP audiences use the results and the decisions they base on them—cannot be easily found. The guidance offered to users varies widely and is often delivered piecemeal, with important details spread across different web pages and reports. Users can obtain NAEP information at three separate websites: the Governing Board site (http://www.nagb.org); the National Center for Education Statistics (NCES) site (http://nces.ed.gov/nationsreportcard/); and a third called “The Nation’s Report Card” (http://www.nationsreportcard.gov). There is some overlap across the three sites in the information available about NAEP, and all have links that take the user from one site to another. But interpretative guidance is uneven across the three, and it can be quite challenging to locate information about the achievement levels (Edley & Koenig, 2017).
Inferences from Various Stakeholders

_Policymakers_

For purposes of this memorandum, we define policymakers as national and state legislators, board and committee members at the federal, state, and district level who make policy and/or recommendations for policy in education, and other individuals who make or influence educational policy (e.g., congressional staffers, lobbyists). These individuals are responsible for policy across educational institutions and have considerable power to influence curriculum, instruction, assessment, teacher professional development, and other factors. They must address information regarding what students know and can do, and whether students are prepared for their next experiences, as policymakers strive to improve the state of American education.

Policymakers use NAEP scores and performance level descriptors for the following purposes:

- making comparisons to other districts, states, and the nation;
- making within-state subgroup comparisons;
- analyzing state achievement trends;
- suggesting changes to state assessments and to aid in defining levels of student performance;
- validating state standards and building the case for educational reform and change in their states (Zenisky, Hambleton, & Sireci, 2009); and
- building arguments for new or amended legislation and for requesting funding related to education (Edley & Koenig, 2017).

NAEP is well-structured in many ways for policymakers, who tend to be most interested in aggregate reports of student performance rather than individual student scores. NAEP is designed to generate comparable results across states and demographic groups. NAEP maintains a scale across years and allows for tracking of trends. However, when policymakers use NAEP to justify changes to state assessments or state performance definitions, build a case for educational reforms, or for requesting funding, they must support those uses based on their own understanding of NAEP and their judgements about NAEP’s suitability for those purposes.

_Educators_

For purposes of this memorandum, we define educators as those persons who work most directly with students. They are responsible for instruction and for implementing curriculum and assessments. Educators include teachers, teachers’ support personnel, content area specialists, academic coaches, etc. We also include school principals in this category, although there is some overlap with policymakers, since principals greatly influence policy within their particular schools.

Because NAEP does not produce results for individual students or at the school level, score interpretations are of limited use for educators. The ALDs and the frameworks, however, may provide considerable useful information. The frameworks indicate the content that students are expected to know in specific subjects at specific grades. The ALDs indicate how students will be categorized based on the level of their knowledge and skill related to that content. The ALDs help educators better understand how student performance is differentiated.
Educators receive their information about NAEP from various sources, including the three main NAEP websites mentioned earlier. They receive much of their information from their state education agency’s website and the media. NCES also supports a NAEP state coordinator in each state who serves as a liaison between the state department of education and the NAEP programs. They are available to assist in the interpretation of NAEP results. We reviewed a sample of state websites as part of preparing this memorandum. We selected websites to reflect either high or low performance on NAEP to highlight any qualitative differences in the information presented to educators.

The three lowest performing states on NAEP 4th and 8th grade reading and mathematics and the three highest performing states based on 2015 results are shown in Table 1. The state Department of Education (DOE) websites and state education agency websites were searched to determine whether and how the states use NAEP data. We specifically searched for information on using NAEP for standard setting purposes.

Table 1. Highest and Lowest Performing States on 2015 NAEP Reading and Mathematics, Grades 4 and 8

<table>
<thead>
<tr>
<th>Subject/Grade</th>
<th>High Performing</th>
<th>Low Performing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mathematics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 4</td>
<td>MA MN NH</td>
<td>AL NM MS</td>
</tr>
<tr>
<td>Grade 8</td>
<td>MA MN NH</td>
<td>AL CA MS</td>
</tr>
<tr>
<td>Reading</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 4</td>
<td>MA NH VT</td>
<td>NM CA AK MS</td>
</tr>
<tr>
<td>Grade 8</td>
<td>NH MA VT</td>
<td>MS NM LA</td>
</tr>
</tbody>
</table>

There were both differences and similarities in how the low and high performing states referred to the available NAEP data. The low performing states provided much less information about participating in NAEP and the purposes of NAEP, in general, compared to the high performing states. High performing states, on the other hand, were more likely to provide details about student performance and participation on NAEP. Many state DOE websites include links to the state NAEP results on the Nation’s Report Card website. Some state websites made a statement that comparisons can be made of how students from different states performed on NAEP, or reference studies that linked state standards to the NAEP standards. However, both low and high performing states provided little information about the explicit uses of the NAEP data for the purposes of creating state level ALDs and informing the determination of cut scores at the state level.

The websites did not include any explicit reference to whether or how NAEP standards may inform state performance standards, or how NAEP data may serve as impact data in state standard settings. The most explicit statement of the connection between state assessment and NAEP was found on the MA DOE website: “…NAEP has taken on a greater prominence under the No Child Left Behind Act and serves to externally confirm results of state assessments, such as the Massachusetts Comprehensive Assessment System (MCAS)” (National Assessment of Educational Progress Frequently Asked Questions, 2017).” The state of Vermont makes

---

another explicit comparison between the structure of its own state science test and the NAEP science assessment standards: “The tests were designed to measure different standards, or frameworks, on separate scoring scales, but both assessments address similar skills and content areas. These assessments provide a way to reference national, state and local science achievement” (Vermont Students Score among Best in the Nation on the National Assessment of Educational Progress, 2016). The state also points out some similarities in the pattern of scores on both the state assessment and NAEP.

Among the state websites studied, most high performing statues reported:

- trends or comparisons of successive cohorts;
- comparison of the percentage of students at or above Proficient on NAEP to the percentage of students at or above Proficient on a state test;
- point-in-time comparisons across states, districts, or population groups (e.g., VT included information showing an increase in the performance of students of low SES);
- performance on subscales (e.g. algebra, vocabulary, etc.)
- rank ordering of states or districts;
- comparisons across population groups to examine performance gaps; and
- comparisons across subject areas.

Lower performing states tended to mention NAEP reports less often. However, we did find some information in the comments of school administrators to the media that NAEP results were used as an indication that the current state education system was in need of reform. For example, in 2013 the then-superintendent of Louisiana, John White, “used the [NAEP state achievement] report to reiterate his push for the Common Core national education standards. ‘The growth this year was moderate. If we want to see something beyond incremental growth, we’ve got to raise our standards, and the Common Core standards is the best way to do that,’ he said” (Bacon-Blood, 2013).

**Researchers and Business Leaders**

For purposes of this memorandum, researchers and business leaders include persons conducting educational research and individuals from private industry with an interest in elementary and secondary student performance. Currently, NAEP data use and interpretation research by these stakeholders may take the following directions (Edley & Koenig, 2017):

- track trends in and compare the performance of successive cohorts,
- make point-in-time comparisons across states and school districts,
- compare the performance of population groups within and across states (performance gaps),
- rank order the performance of states and compare state to national performance;
- compare performance across tested subject areas,
- examine relationships among student performance and selected student/school/family variables, and
• compare states’ standards for proficient performance in reading and mathematics by placing them on a common scale defined by NAEP scores (“mapping studies”).

Beginning with NAEP results from 2003, NCES conducted a series of studies that mapped each state’s grade 4 and 8 reading and mathematics proficiency levels to the NAEP scale. This mapping was designed as a mechanism to evaluate the extent to which state standards reflected the same rigor as NAEP standards, and it was used as a policy lever to encourage states to set challenging standards for their students (Edley et al., 2017). In the mapping study report by Bandeira de Mello, Bohrnstedt, Blankenship, & Sherman (2015), the NAEP score that corresponds to a state’s standard (i.e., the NAEP scale equivalent score) is determined by a direct application of equipercentile mapping. For a given subject and grade, the percentage of students reported in the state assessment to be meeting the standard in each NAEP school is matched to the point on the NAEP achievement scale corresponding to that percentage. The percentage of students passing the state standard was mapped onto the NAEP scores. The results are then aggregated over all of the NAEP schools in a state to provide an estimate of the NAEP scale equivalent of the state’s threshold for its standard (Bandeira de Mello et al., 2015).

Peterson and Ackerman (2015) took a different approach to the comparison of state achievement scores and NAEP scores. They calculated the difference between the percentage of students considered “proficient” by both the state and NAEP assessments. The magnitude of the difference was considered to indicate how rigorous the state standards are as compared with NAEP standards.

These examples indicate that some researchers and policymakers do consider NAEP achievement levels to be a standard that states should strive toward. At the same time, some researchers caution against using NAEP as an infallible measure of state educational achievement due to fundamental differences between the state and NAEP frameworks and standards (e.g., Ho & Haertel, 2007). It is important to remember that determining the score equivalency between NAEP scale and state scale does not say anything about the equivalency or lack thereof in knowledge and skills associated with the score. The NAEP and state assessments may or may not measure the same knowledge and skills. An alignment study would need to be conducted to assess the extent to which the two assessments measured the same construct.

Many studies focused on validity evidence based on relationships with external variables, that is, setting benchmarks on NAEP that are related to concurrent or future performance on measures external to NAEP. Examples are academic preparedness for college; international tests; state tests and their alignment with NAEP (Edley et al., 2017). The studies indicate that there is considerable correspondence between the percentages of students at NAEP achievement levels and the percentages on other assessments (Gattis et al., 2016; Jia et al., 2014; Lim & Sireci, 2017; Neidorf, Binkley, Gattis, & Nohara, 2006; Phillips, 2014a, 2014b; Poland & Plevyak, 2015; Provasknik, Lin, Darling, & Dodson, 2013). These studies show that the NAEP achievement-level results (the percentage of students at the advanced level) are generally consistent with the percentage of U.S. students scoring at the reading and mathematics benchmarks on the Programme for International Student Assessment (PISA), the mathematics benchmarks on Trends in International Mathematics and Science Study (TIMSS), and at the higher levels for College Board Advanced Placement (AP) exams. For example, a report by Fields (2014) states that the content of the 12th grade NAEP reading and mathematics assessments was found to be similar to widely recognized tests used for college admission and placement. A linking study by Moran, Freund, & Oranje (2012) determined that there is a higher correlation between NAEP and SAT mathematics scores than between NAEP and SAT reading.
scores. The SAT reading benchmark, however, was closer to the NAEP Proficient score than the SAT math benchmark. Several studies investigated the relationship between NAEP Proficient and college and career readiness (Moran, Oranje, & Freund, n.d.; Schneider, Kitmitto, Muhusani, & Zhu, 2015), but the relationship was found to be fairly weak. Additional research in this area was proposed.

During the August 2016 Governing Board quarterly meeting, researchers provided the following recommendations regarding the use of NAEP data.

- Panelists urged the Governing Board to enable linkages from NAEP data to state-level or national-level to conduct research about the long-term effects of educational policies.
- All panelists agreed that while NAEP data describe trends in student achievement, the data do not support conclusions about the reasons for these trends. Additional research is needed to discover factors that can improve schools and student learning.
- It was suggested that the NAEP data be used to compare the performance of districts with similar demographic characteristics, such as poverty levels. NAEP data may be used to guide best practices on what works in the improvement of educational achievement.

The Media

While academic and research articles provide scientific, well-reasoned rationales for or against the specific interpretations of NAEP, articles by the media present a different side. They tell the story of those who are trying to use information under real-life conditions from the assessments that the academics are studying, and the real-world challenges and issues experienced by practitioners in the field.

Articles in publications like Education Week illustrate that there is a large degree of confusion accompanying the application and interpretation of NAEP standards. While many researchers and even state officials may assume the debate about the application of NAEP standards is resolved, magazine and newspaper articles question whether it is appropriate for states to incorporate NAEP standards into the standards of the state, and what the appropriate uses for NAEP scores are in general.

One point of argument is lack of clarity on the meaning of “proficient” and the application of that meaning to state standards. Not all media representatives consistently clarify for the public that NAEP Proficient is not grade-level proficiency and that NAEP Proficient is intended to be an aspirational standard. What makes this matter more complicated is that under the No Child Left Behind Act (NCLB), states had to create achievement levels that were grade-specific and most states chose to adopt the ALD title of “Proficient.” Reconciling these sets of standards causes additional conflict and confusion when states are trying to create their achievement levels and communicate them to the public. One suggestion to make the situation more understandable is for policymakers to explain to the stakeholders “what are good goals for educational purposes compared to what is appropriate for accountability when establishing cut scores on their state assessments” (Hull, 2008), why they may be different, and which performance levels are more appropriate for each specific purpose.
Many researchers are concerned that information from NAEP gets misinterpreted by the media and politicians, sometimes to serve the interests of specific groups. Various misinterpretations of NAEP results are frequently used by the politicians and media, giving rise to the term “misnaepery” (Sawchuk, 2013). One prominent example of this inappropriate interpretation includes tying an increase in state NAEP scores to some specific policy or intervention implemented by the state, and a decrease – to a policy that was proposed by an organization, but then not implemented. In practice, it is very challenging to make these causal connections. Organizations that are using NAEP scores to bolster claims about the effects of a specific policy are likely not interpreting the NAEP scores correctly (Chingos & Blagg, 2015).

A number of misinterpretations come from the misunderstanding of NAEP’s definition of “proficient”, with some reporters claiming that being below proficient means being “below grade level.” Yet another source of confusion comes from comparing state assessment scores with NAEP scores and arriving at opposing conclusions. Comparing the achievement of different student population groups is often fraught with misinterpretations as well (e.g., treating the NAEP achievement scale as continuous between grades and comparing achievement of one population at a higher grade to the achievement of another population at a lower grade).

At least in part, these misinterpretations arise from a lack of readily available or accessible information on how the NAEP scores should be interpreted, what the appropriate uses of these scores are, and what conclusions are appropriate to make. Educational researchers call for using caution in deciphering which claims are appropriate, and discouraging the propagation of false claims about NAEP data interpretation (Polikoff, 2015a, 2015b).

**The General Public**

The general public may not have sufficient knowledge and training to deeply understand the intent and the meaning of state or national assessments, and may have a difficult time interpreting and critically evaluating information coming from various, often conflicting, sources. The media may make the situation in education appear more critical or negative than it really is. For example, if a state performs as one of the best on NAEP, but there is no growth in scores, the general public may see headlines like “Public education test results are dismal. Schools are failing NH children” (Levell, 2016). In addition, as mentioned earlier, the information provided by the media may not be completely objective, and score interpretations may be promoting a specific political agenda.

There is some confusion among the general public regarding why their state may have high scores on the state assessments, but low scores on NAEP (Weiss, 2016; Dillon, 2005). This may occur if the state set standards lower than NAEP standards, or if the state simply has different content standards. There may also be conflicting information on exactly how the state standards compare to NAEP standards; this may cause one study to claim that a state has low standards, and another study – that the state is either lagging behind others, or low on scores from some other perspective. A study by Achieve⁴ describes several NAEP objectives at grade 4 contrasted with the grade those same objectives are introduced in several states’ standards documents. The objective “Use simple ratios to describe problem situations,” is typically introduced in grade 6 in many states. Discrepancies like this add complexity to potential comparisons between NAEP results and state testing results.

---

One potential goal would be for the general public to be able to use state and national assessments to make decisions about whether children are getting the best education in their particular state. It is likely impossible to make such inferences at the school or even classroom level from state and national assessments. The media, however, may make it sound like those conclusions are appropriate and necessary. The same article by Levell (2016) that proclaimed the failure of New Hampshire public education, for example, suggests that, based on the fact that there was little to no growth in the student scores on state assessments or NAEP, the parents should “[e]ngage your local school board and question why they are using College and Career Readiness Standards and tests that are not providing a better education for our children;” consider a transfer to a charter or private school; or refuse to have their child take a state assessment. It may be helpful for the general public to have access to a source of clear, easy to understand, reliable information on the kinds of inferences that can legitimately be made from state and national assessments.

**Approaching Validation of the NAEP Performance Levels Using a Validity Argument**

A strong validity argument relies upon a foundation of thorough and specific definitions of the various purposes of the assessment. These purposes are typically illustrated via a *Theory of Action* (TOA) document or graphic. The TOA indicates the intended uses and expected impact of the assessment system. As depicted in Figure 1, the TOA can inform testable claims related to the interpretation of test scores. These testable claims represent the *interpretive argument*. Every use or interpretation of an assessment score relies on meeting specific claims and the various assumptions that justify them. The evidence supporting those assumptions represents the *validity argument*. The NAEP assessments represent a large number of potential interpretations/uses for test scores.

![Figure 1. Relationships among theory of action (TOA), interpretive argument, and validity argument.](image)

The Governing Board’s Strategic Vision indicates that NAEP results should inform stakeholders “about what America’s students know and can do in various subject areas and compare achievement data over time and among student demographic groups” (p. 1). The ALDs provide context for that goal by helping stakeholders interpret student performance in the various subject areas. Estimates of the proportions of students who would be classified as below Basic,
Basic, Proficient, or Advanced for each state, for select large school districts, and for demographic groups of students within them are reported. Reports are generated based on the performance of representative groups of students within those states and districts.

The subject matter content tested by NAEP and the ways student mastery of that content are operationalized in the achievement levels are described in the frameworks documents. These documents are vital to the TOA and to the interpretive argument. They describe what is tested on each of the NAEP subject tests and help us differentiate student performance into meaningful categories. If we were to construct a chain of logic, as is typically done in a TOA, the following assertions might be included.

The subject area content included in the frameworks represents important key knowledge, skills, and concepts students should know at the indicated grade level.

1. The ALDs differentiate important differences in students’ mastery of the content included in the frameworks.
2. NAEP assessments allow for strong estimates regarding the proportions of students scoring in each of the performance categories.
3. Score reports, or report cards, can be referenced to the frameworks and ALDs to interpret what students within a given state or large district know and can do.
4. Comparisons across states, large districts, and demographic groups allow stakeholders to identify gaps in terms of what students know and can do.
5. Stakeholders use NAEP performance information to better understand student achievement in their efforts to improve the education of American students.

The next step toward constructing the validity argument is to use the chain of logic from the TOA to describe how inferences from test scores are used by stakeholders in the process of achieving the goals of the testing program. When we consider the interpretive argument, we are forced to imagine the role of the various stakeholders. As an example, if we were to assume the role of a state education agency stakeholder, we might interpret NAEP results in the following ways, among others.

1. My state NAEP scores provide a snapshot of student performance for the current year’s students’ performance in the tested subjects.
2. My NAEP scores represent student achievement for the academic content the students are expected to learn, as described in the NAEP framework for each subject.
3. My state scores can be directly compared to other states and those comparisons will tell me if my state is preparing students as well as other states.
4. Demographic groups of students can be compared to each other for my state, and those comparisons give me information about performance gaps among those groups.
5. By comparing demographic group performance across states, I can determine if my state’s performance gaps are larger or smaller than the gaps in other states.
6. The proportions of students from my state in each performance level are in those levels because of differences in their preparation related to knowledge, skills, and abilities as described in the ALDs.
7. I can directly compare my NAEP results this year to prior year’s results to determine if students in my state are improving, declining, or staying at about the same level in the tested subjects and grades.

The next step in the process of building a validity argument would be to support the inferences described above through a claims and evidence structure. The claims are usually written as a series of “if…then” statements. The claims support the specific inference described in the interpretive argument. If we take #6 from the list of inferences above “The proportions of students from my state in each performance level are in those levels because of differences in their preparation related to knowledge, skills, and abilities as described in the ALDs,” the claims might include the following.

1. If NAEP test items are designed to differentiate the skills associated with the knowledge, skills, and abilities described in the ALDs, then NAEP scores may relate directly to the ALDs.
2. If NAEP content is sufficiently similar to the content educators teach in schools, then NAEP scores may reflect students’ preparation in schools.
3. If student preparation in schools improves, then NAEP scores should also improve.

There are other claims that might be needed to support this inference, but these provide an example of the structure of the validity argument. The claims are then arranged in a structure or graphic that indicates their interconnected nature and dependencies. Failure to support one claim may undermine all subsequent claims that depend on it. For example, the frameworks define the NAEP assessment content. If that content were substantively different from the content taught in schools within a state, NAEP’s validity for determining if the students were improving from year to year would be compromised. The students might be improving greatly on content extraneous to NAEP. All inferences related to subgroup performance or subgroup gains would also be undermined. Comparisons to other states, with content similar to that tested on NAEP, would also be undermined.

For the final step, one would simply summarize the evidence supporting each of the claims and determine if the claim is supported, not supported, or if there is insufficient evidence to draw a conclusion. For many claims, previously collected evidence can simply be referenced. For other claims, new investigations may be needed or updates to existing research may be required to account for changes in the American education system, contextual variables that threaten validity, or other factors.

The validity argument might be structured in any number of ways, but a simple approach is to generate tables that include claims, assumptions, evidence, and support. Table 2 provides one example of how a portion of a NAEP validity argument related to the achievement levels might look. The claims are abbreviated from the list of “if…then” statements above and are leftmost in the table. The next column contains the assumptions that underlie this claim. The third column lists evidence that might be used to support the assumptions. The final column is for a summary judgement regarding whether the evidence is supportive (S), non-supportive or counter to the assumption (N), or inconclusive (I). Mock values for this final column are provided in Table 2 to illustrate one way that the validity argument might be constructed. These values do not represent an evaluation of the evidence available.
### Table 2. Test Design Claims, Assumptions, and Evidence

<table>
<thead>
<tr>
<th>Claim</th>
<th>Assumptions</th>
<th>Evidence</th>
<th>Summary Judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Items Differentiate NAEP Achievement Levels</strong></td>
<td>Items were written to reflect NAEP achievement levels.</td>
<td>Item writing guidelines, instructions, and documentation reflect achievement levels.</td>
<td>S</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Item coding in metadata is linked to achievement levels</td>
<td>S</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Each of the achievement levels is well represented in the item pool for all content categories.</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ALD classification accuracy is acceptably high.</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Item and test statistics support classification of students.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Metadata supports classification (e.g., the most difficult items reflect the descriptions in the higher achievement levels).</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Documentation from standards-setting activities indicate appropriate processes were followed.</td>
<td>S</td>
</tr>
<tr>
<td><strong>2. NAEP tests the content taught in schools</strong></td>
<td>Content from NAEP Frameworks largely coincide with state academic standards.</td>
<td>Alignment studies indicate substantial correspondence of content.</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The depth described in the NAEP ALDs is similar to the depth described in state performance level descriptors.</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Schools teach the main categories of content described by the Frameworks</td>
<td>I</td>
</tr>
<tr>
<td><strong>3. Improvements in student preparation are reflected on NAEP</strong></td>
<td>NAEP results are sensitive to major changes in educational practice.</td>
<td>Analysis of trend data tracks the timing of major state reform efforts.</td>
<td>S</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NAEP gains/losses are reflected in similar measures of student performance.</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Comparisons of gains scores on NAEP are consistent with gains on statewide assessments, ACT, SAT, etc.</td>
<td></td>
</tr>
</tbody>
</table>
Contextual Factors that Represent Challenges for Constructing a Validity Argument for NAEP Achievement Levels

One of the most challenging aspects of validation for NAEP ALDs is the context in which NAEP scores are interpreted. The ALDs differentiate students into “Proficient” versus “not-Proficient” categories, and those labels are common with federal requirements for state assessments. It is common for the media to compare state results to NAEP results. When states declare a larger proportion of students to be proficient than NAEP, that finding is often taken as evidence that the state’s standards are less rigorous. When NAEP reports that a substantive proportion of students score lower than proficient, those results can be characterized as indicative that students are not on grade level, or that they are unprepared for the next stage in their educational experiences.

These inferences are not supported by NAEP’s official documentation, but they are so common that it might be beneficial to consider them when constructing a validity argument. It may be beneficial to characterize the NAEP achievement levels in the context of other common metrics or common understanding of terms. For example, there are multiple indicators of readiness for college (e.g., ACT and SAT benchmarks, specific high school course grades, placement tests, etc.). Many of these indicators have been validated based on outcome criterion (e.g., college course grades, advancement from year 1 to year 2 in college, or attainment of a degree). Providing context related to the NAEP achievement levels that reference similar information may help with interpretation. NAEP is not designed as a college entrance exam, nor as a specific indicator of college readiness. However, indicating that students who score in a particular category tend to also meet other indicators of college readiness could help stakeholders make more sense of their scores.

Another key way that the achievement levels are used by educators is as a guide for what content students are expected to learn and to what degree they are expected to learn that content. The frameworks and the achievement levels provide guidance on expectations for educators, especially in subjects other than mathematics and reading/English language arts, where there may not be clear state standards documents. The frameworks may be used less for mathematics and reading because all states were required to adopt standards for those subjects by federal mandate under the No Child Left Behind Act. Later, most states adopted the Common Core State Standards (CCSS), either in their entirety or with minor editing. These CCSS now serve to guide much of the content taught in American schools. States typically individually worked to characterize performance in relation to the CCSS, so despite common content standards, performance standards vary substantially by state. The NAEP Frameworks and achievement levels are secondary indicators of what students should know and be able to do. If there are important differences between the two standards documents, it could undermine the validity of NAEP scores. If performance is categorized differently by the state for the CCSS than for NAEP, it becomes a challenge for educators to reconcile the differences. Depending on how the states define “Proficient” in reference to the CCSS, educators may not be striving toward “Proficient” as defined by the NAEP achievement levels even if the content of the state assessment and NAEP are largely the same.

There are other contextual factors that should be considered related to the NAEP achievement levels. These factors represent a challenge when drawing inferences from NAEP results and may foster misunderstandings and misuses of data. Their impact can be attenuated by clear guidance regarding the inferences that are supported and those that are not.

5 See http://www.corestandards.org/.
Using NAEP Achievement Levels to Inform Statewide Testing Standards

One way that NAEP achievement levels have been used by state policymakers is to inform cut scores during standards setting for their statewide achievement tests. States are required to test students in reading/English language arts (ELA) and mathematics in grades 3-8 and high school under the Every Student Succeeds Act (ESSA). Many states also have statewide tests for science and social studies in selected grades. States are required to report results in terms of the proportion of students scoring at the “Proficient” level or above. The level of reporting and the use of the common performance category “Proficient” leads many stakeholders to make comparisons between statewide testing results and NAEP results. States may be criticized if a much greater proportion of students are classified as proficient in grade 4 mathematics on the statewide test than are classified as proficient on NAEP. One of the ways that some states avoid this criticism is to include NAEP achievement levels as part of their standards setting procedures.

While there are several ways that states might include NAEP results in their standards setting, we will consider two here. The first is to use NAEP results as impact data. This use of NAEP may or may not impact cut scores set for state assessments. NAEP results are often used as part of a set of impact data—so the proportions of students in each achievement level on NAEP are considered in conjunction with other information (e.g., the proportion meeting college benchmarks, the proportion in each of the state’s reporting categories for a prior assessment, etc.) prior to assigning final cut scores. This typically occurs after standards setting panelists have completed at least one round of assigning cut scores. Impact data is used as a “reality check” to determine if the state cut scores will create controversy in light of other information.

Using NAEP achievement levels to generate impact data requires little in the way of validity evidence, as long as the standards setting facilitators make clear that no direct relationship is expected between NAEP and state assessment results. If, however, the facilitators do not make clear that NAEP achievement levels do not imply grade level performance, college readiness, or other inferences, this impact data can have a much more significant impact on the state’s cut scores. If such inferences were intended, a great deal of validity evidence would be needed to support them. Some standards setters guard against making sweeping changes during later rounds of the process, when impact data reviewed, by placing limits on how far the cut scores can be moved at each stage. This prevents panelists from basing their cut scores on impact data to the exclusion of the performance level descriptors and/or test items.

On the other end of the spectrum, states could create cut scores for their assessments that mirror NAEP achievement levels. This could be accomplished through an equipercentile process without using panelists. It is more likely that the equipercentile solution is presented to panelists as a starting point for standards setting. Then, based on the state’s performance level descriptors and/or items, panelists might move the cut scores in one direction or the other to better align with the state’s overall assessment system. Limits might be placed on how far the cut scores could deviate to ensure that the proportions of students in each classification category were similar to NAEP. This process would assure that state assessments had similar rigor to NAEP and would allow for more coherent comparisons between the state system and NAEP.

The validity evidence needed to support using NAEP achievement levels in this way would be much more stringent. First, the state would need to ensure that the content of the two tests were sufficiently similar to support consistent cut scores. This would likely require an alignment study. Then, the state would need to establish that the performance level descriptors for the statewide...
assessment and for NAEP captured much the same kinds of performance and referenced similar differentiators for each performance category. If not, students might exhibit qualitatively different skills on the assessments, despite scoring similarly.

Evaluating NAEP’s Achievement Levels for an Evolving Educational Landscape

NAEP tests students in specified grades in several subjects. Reading and mathematics are tested every other year, while other subjects are tested less often. NAEP’s achievement levels for math and reading were established in the early 1990s, while achievement levels for some of the other subjects (e.g., writing, science) have been set or revised more recently. It is important to consider the claims and assumptions that led to the creation of NAEP achievement levels and to verify that those claims and assumptions continue to be relevant and supported as education in America evolves. It is important to verify that NAEP continues to measure the most important content for the tested subjects, that those subjects are the most relevant for stakeholders, and that the knowledge, skills, and abilities described in the achievement levels still represent the most important differentiators for student achievement. A strong validity argument is not static, but routinely tests its claims and assumptions as the inferences stakeholders draw from test information change.

Summary: Steps Toward Developing a Validity Framework for NAEP Achievement Levels

The most important step toward validation of the NAEP achievement levels is to explicitly state the inferences that are expected to be made. These inferences will guide the creation of the specific validity claims, which in turn will help the Governing Board organize and present evidence to support the use of the Achievement Levels for their designated purposes. This priority is in line with the Governing Board’s Strategic Vision and is explicit in its response to the achievement levels evaluation (National Assessment Governing Board, 2017).

Once the inferences are made explicit, the next step in the validation process is to investigate the utility of the Achievement Levels for their intended purposes. We know that one of those purposes is to help define what students know and can do within the tested subjects. The ALDs describe student performance within specific ranges on the scale. Users of NAEP data are provided with the proportions of students expected to be at each performance level, which they interpret in conjunction with the ALDs. It would be beneficial to sample from these interpretations to ascertain if the information provided is meeting the needs of key stakeholders, and to determine if those stakeholders are making unsupported interpretations from the data.

This process will provide key input into the next step in establishing a validity framework, the creation of an interpretive guide for NAEP achievement levels. Such a guide would indicate the key inferences stakeholders are expected to make, caveats and limitations on those interpretations, and warnings about common potential misinterpretations or misuses of the NAEP Achievement Levels or achievement level data. The interpretive guide should not be limited to achievement levels, but also include information on the use of scale scores, comparisons across jurisdictions (e.g. states or large districts), and it should describe when it is most appropriate to use achievement levels versus scale scores.

Once the interpretive guide is complete, it can be used to guide the remainder of the validity argument. For example, if the interpretive guide characterizes the content in the Achievement Level for fourth grade Science at Basic as the content that the typical student scoring at that level has mastered, validity evidence would be needed to support that statement. The content described for the Basic level of fourth grade science might be compared with the content of the
NAEP test items that best discriminate within the Basic range of the scale. If the item content essentially matched the content described in the ALD, that finding would represent support for the interpretation. There is, of course, other evidence that might also be used to support such an interpretation. The inference would be considered valid if the preponderance of this evidence was supportive and no evidence directly contradicted the inference.

This process would be repeated for each of the inferences described in the interpretive guide until all the inferences were addressed to the satisfaction of assessment validity experts, several of whom serve on the Governing Board. For many of the intended inferences, it will be possible to simply reference research that has already been completed. For other inferences, it may be necessary to conduct additional research in order to bring appropriate evidence to bear. If any of the inferences is unsupported by evidence or if the evidence that is available is negative, either the interpretation must be altered or the test information bolstered in some way. The evidence included in the validity argument may need to be revised or updated any time the NAEP assessments are revised or altered, any time there is a significant shift in the national educational landscape, and when there are concerns that the evidence is so dated that it may no longer be applicable.
References


Communication and Interpretation of Achievement Levels

At the November 2018 Governing Board meeting, COSDAM and the Reporting and Dissemination Committee will hold a joint meeting to discuss the two committees’ work on achievement levels for the National Assessment of Educational Progress (NAEP).

**Background**

From 2014 to 2016, the National Academies of Sciences, Engineering, and Medicine evaluated the NAEP achievement levels in mathematics and reading, which are the responsibility of the Governing Board. In their evaluation, the National Academies noted eight common uses of NAEP achievement levels, specifically:

- Trends or comparisons of successive cohorts, e.g., the percentage of students at or above Proficient in reading has increased over time;
- Comparison to a state assessment;
- Point-in-time comparisons across states, districts, or population groups, e.g., more students in state A who are at or above Proficient in reading compared to state B;
- Rank ordering states or districts;
- Comparison across population groups to examine performance gaps;
- Comparison across subject areas, e.g., more students perform at or above Proficient on mathematics than in reading;
- Comparison of before and after an action or policy implementation; and
- Relationships among achievement results and contextual data.

The evaluation recognized the usefulness and value of the achievement levels but made several important recommendations, most of which focus on the work of COSDAM as well as two that also address the work of the R&D Committee:

**RECOMMENDATION 5:** Research is needed to articulate the intended interpretations and uses of the achievement levels and collect validity evidence to support these interpretations and uses. In addition, research to identify the actual interpretations and uses commonly made by NAEP’s various audiences and evaluate the validity of each of them. This information should be communicated to users with clear guidance on substantiated and unsubstantiated interpretations.

**RECOMMENDATION 6:** Guidance is needed to help users determine inferences that are best made with achievement levels and those best made with scale score statistics. Such guidance should be incorporated in every report that includes achievement levels.

Since the release of these recommendations in November 2016, Governing Board staff and COSDAM members have started working to fulfill these recommendations. The draft revision of
the Board policy on developing student achievement levels (planned for full Board action in November 2018) establishes an

“interpretative guide [which] shall accompany NAEP reports, including specific examples of appropriate and inappropriate interpretations and uses of the results” (Principle 3h).

This guide is intended for inclusion on the Nation’s Report Card website and on specific report card webpages. The guide will target stakeholders, such as media, policy advocates, members of the general public, educators, and policymakers. These groups may be familiar with both NAEP and achievement levels, but their understanding, interpretation, and use of achievement levels could be informed and improved with guidance from the Governing Board.

The Reporting and Dissemination Committee will collaborate with COSDAM on the development of this interpretative guide. The overarching question of the joint meeting will focus on the general approach the interpretative guide should take. This joint meeting also will elicit feedback on several specific features of the guide:

(1) the scope—what should be covered and what should not;
(2) the content—uses of achievement levels, value and usefulness of achievement levels;
(3) the language—non-technical, accessible; and
(4) the delivery—how the guide will be included with report cards.

If there is time, the conversation may extend to initial discussions of a statement on both the uses and usefulness of NAEP generally, not only of achievement levels specifically.

In addition, the committees should deliberate together on how to engage stakeholders on improving their use and interpretation of NAEP and achievement levels beyond the interpretative guide.
Strategic Vision Activities Led by COSDAM

During the November 2016 Board meeting, a Strategic Vision was formally adopted to guide the Board’s work over the next several years. For each activity led by COSDAM, information is provided below to describe the current status and recent work, planned next steps, and the ultimate desired outcomes. Please note that many of the Strategic Vision activities require collaboration across committees and with NCES, but the specific opportunities for collaboration are not explicitly referenced in the table below. In addition, the activities that include contributions from COSDAM but are primarily assigned to another standing committee (e.g., framework update processes) or ad hoc committee (i.e., exploring new approaches to postsecondary preparedness) also have not been included below.

The Governing Board’s Assistant Director for Psychometrics, Sharyn Rosenberg, will answer any questions that COSDAM members have about ongoing or planned activities.

<table>
<thead>
<tr>
<th>Strategic Vision Activity</th>
<th>Current Status and Recent Work</th>
<th>Planned Next Steps</th>
<th>Desired Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>SV #2: Increase opportunities to connect NAEP to administrative data and state, national, and international student assessments</td>
<td>Ongoing linking studies include: national NAEP-ACT linking study; longitudinal studies at grade 12 in MA, MI, TN; longitudinal studies at grade 8 in NC, TN; NAEP-TIMSS linking study; NAEP-HSLS linking study; NAEP Validity Studies (NVS) studies. Informational update on current studies was provided in the March 2018 COSDAM materials. As of October 2018, analyses are currently underway for the national NAEP-ACT linking study, with presentation to COSDAM tentatively planned for March 2019.</td>
<td>Complete ongoing studies. Decide what new studies to take on. Decide how to use and report existing and future results. Complete additional studies.</td>
<td>NAEP scale scores and achievement levels may be reported and are better understood in terms of how they relate to other important indicators of interest (i.e., other assessments and milestones).</td>
</tr>
<tr>
<td>Strategic Vision Activity</td>
<td>Current Status and Recent Work</td>
<td>Planned Next Steps</td>
<td>Desired Outcome</td>
</tr>
<tr>
<td>----------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>SV #3: Expand the availability, utility, and use of NAEP resources, in part by creating new resources to inform education policy and practice</td>
<td><em>Research when and how NAEP results are currently used (both appropriately and inappropriately) by researchers, think tanks, and local, state and national education leaders, policymakers, business leaders, and others, with the intent to support the appropriate use of NAEP results (COSDAM with R&amp;D and ADC)</em></td>
<td>Use research to draft short document of intended and appropriate uses for COSDAM discussion (March 2019)</td>
<td>Board adopts formal statement or policy about intended uses of NAEP. The goal is to increase appropriate uses and decrease inappropriate uses (in conjunction with dissemination activities to promote awareness of the policy statement)</td>
</tr>
<tr>
<td></td>
<td><em>Develop a statement of the intended and unintended uses of NAEP data using an anticipated NAEP Validity Studies Panel (NVS) paper and the Governing Board’s research as a resource (COSDAM with NCES)</em></td>
<td>NCES produces documentation of validity evidence for intended uses of NAEP scale scores</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Disseminate information on technical best practices and NAEP methodologies, such as training item writers and setting achievement levels</em></td>
<td>Governing Board produces documentation of validity evidence for intended uses of NAEP achievement levels</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ina Mullis of the NVS panel spoke with COSDAM at the March 2017 Board meeting and is working on a white paper about the history and uses of NAEP</td>
<td>A joint discussion of COSDAM and the Reporting &amp; Dissemination Committee was planned for November 2018 but has been postponed to March 2019 to allow time for focused discussion on achievement levels instead</td>
<td></td>
</tr>
<tr>
<td></td>
<td>During the August 2018 Board meeting, COSDAM discussed how to use information from an ongoing study to inform a policy statement on intended and appropriate uses of NAEP</td>
<td>This idea was generated during the August 2017 COSDAM discussion of the Strategic Vision activities</td>
<td>Stakeholders benefit from NAEP technical expertise</td>
</tr>
<tr>
<td></td>
<td>A joint discussion of COSDAM and the Reporting &amp; Dissemination Committee was planned for November 2018 but has been postponed to March 2019 to allow time for focused discussion on achievement levels instead</td>
<td>Work with NCES and R&amp;D to refine list of technical topics for dissemination efforts</td>
<td></td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Strategic Vision Activity</th>
<th>Current Status and Recent Work</th>
<th>Planned Next Steps</th>
<th>Desired Outcome</th>
</tr>
</thead>
</table>
| SV# 5: Develop new approaches to update NAEP subject area frameworks to support the Board’s responsibility to measure evolving expectations for students, while maintaining rigorous methods that support reporting student achievement trends | *Consider new approaches to creating and updating the achievement level descriptors and update the Board policy on achievement levels*  
Input for the policy revision was provided through a panel of standard setting experts, a literature review on considerations for creating and updating achievement level descriptors (ALDs), and a technical memo on developing a validity argument for the NAEP achievement levels (early 2018)  
COSDAM discussed the policy revision during the May and March 2018 Board meetings  
Full Board discussed the draft revised policy during the August 2018 Board meeting  
Public comment was sought from August 30 – October 15, 2018; Board calls to discuss the comments took place in October  
Additional discussion of the draft revised policy will take place during the upcoming November 2018 Board meeting | Board action on revised policy statement (planned for November 2018)                                                                                                                                                                                                             | Board has updated policy on achievement levels that meets current best practices in standard setting and is useful for guiding the Board’s achievement levels setting work |
<table>
<thead>
<tr>
<th>Strategic Vision Activity</th>
<th>Current Status and Recent Work</th>
<th>Planned Next Steps</th>
<th>Desired Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>SV# 7: Research policy and technical implications related to the future of NAEP Long-Term Trend assessments in reading and mathematics</td>
<td>White papers commissioned, symposium held in Washington, DC (March 2017), and follow-up event held at American Educational Research Association (AERA) annual conference (April 2017)</td>
<td>Per the discussion and next steps at the March 2018 Executive Committee meeting, COSDAM will discuss design considerations for the next administration of LTT. Pending the outcome of discussions in the Executive Committee meeting in November, additional information about design considerations will be shared with COSDAM at the March 2019 meeting.</td>
<td>Determine whether changes to the NAEP LTT schedule, design and administration are needed (led by Executive Committee and NCES)</td>
</tr>
<tr>
<td>Support development and publication of multiple papers exploring policy and technical issues related to NAEP Long-Term Trend. In addition to the papers, support symposia to engage researchers and policymakers to provide stakeholder input into the Board’s recommendation</td>
<td>Full Board and Executive Committee discussions (March, May, and August 2017) and webinar on secure LTT items and p-values from 2012 administration (October 2017)</td>
<td>The NAEP budget in Fiscal Year 2019 has been increased by $2 million with a goal of moving up the next administration of LTT (Discussion in November 2018 Executive Committee meeting)</td>
<td></td>
</tr>
<tr>
<td>SV# 9: Develop policy approaches to revise the NAEP assessment subjects and schedule based on the nation’s evolving needs, the Board’s priorities, and NAEP funding</td>
<td>COSDAM presentation and discussion on initial considerations for combining assessments</td>
<td>Board action on the NAEP Assessment Schedule tentatively scheduled for March 2019</td>
<td>Determine whether new assessment schedule should include any consolidated frameworks or coordinated administrations</td>
</tr>
<tr>
<td>Pending outcomes of stakeholder input (ADC activity), evaluate the technical implications of combining assessments, including the impact on scaling and trends</td>
<td>During the past year, there have been several full Board presentations and discussions on the assessment schedule</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Initial draft schedule and budget to be discussed in November 2018</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strategic Vision Activity</td>
<td>Current Status and Recent Work</td>
<td>Planned Next Steps</td>
<td>Desired Outcome</td>
</tr>
<tr>
<td>---------------------------</td>
<td>-------------------------------</td>
<td>-------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>SV# 10: Develop new approaches to measure the complex skills required for transition to postsecondary education and career</td>
<td>Several studies are ongoing (see activities under SV# 2)</td>
<td>Decide whether additional research should be pursued at grade 8 to learn more about the percentage of students “on track” to being academically prepared for college by the end of high school or whether additional research should be conducted with more recent administrations of NAEP and other tests</td>
<td>Statements about using NAEP as an indicator of academic preparedness for college continue to be defensible and to have appropriate validity evidence</td>
</tr>
<tr>
<td></td>
<td>Per COSDAM discussion at August 2017 meeting, additional studies are on hold until at least November 2018 pending Board decision on how to move forward with findings from Ad hoc Committee on Measures of Postsecondary Preparedness</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Continue research to gather validity evidence for using 12th grade NAEP reading and math results to estimate the percentage of grade 12 students academically prepared for college</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# National Assessment Governing Board
## Reporting and Dissemination Committee
### Friday, November 16, 2018
10:30 am – 12:30 pm

## AGENDA

<table>
<thead>
<tr>
<th>Time</th>
<th>Item</th>
</tr>
</thead>
</table>
| 10:30 – 10:35 am | Welcome and Information Items  
  *Rebecca Gagnon, Chair* |
| 10:35 – 10:50 am | Developing Messaging for Governing Board Members  
  *Laura LoGerfo, Assistant Director for Reporting and Analysis* |
| 10:50 – 11:05 am | **ACTION:** Release Plan for 2018 Technology and Engineering Literacy Assessment  
  *Stephaan Harris, Assistant Director for Communications* |
| 11:05 – 11:30 am | What NCES Data Could Work with NAEP Data (SV #2)  
  *Robert Finnegan, ETS*  
  *Dan McGrath, National Center for Education Statistics* |
| 11:30 am – 12:30 pm | Joint Meeting with Committee on Standards, Design and Measurement (COSDAM): Communication and Interpretation of NAEP Achievement Levels (SV #3)  
  *Rebecca Gagnon*  
  *Andrew Ho, Chair, COSDAM* |

<table>
<thead>
<tr>
<th>Attachment</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
</table>
## NAEP Reports for Release in Fall 2018

<table>
<thead>
<tr>
<th>Title</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015 Student Questionnaires: Classroom Instruction for Mathematics Reading and Science</td>
<td>October 2018</td>
</tr>
<tr>
<td>2015 National Indian Education Study: A Closer Look</td>
<td>October 2018</td>
</tr>
<tr>
<td>Findings and Recommendations from the National Assessment of Educational Progress (NAEP) 2017 Pilot Study of the Middle School Transcript Study (MSTS)</td>
<td>October 2018</td>
</tr>
<tr>
<td>Responsibility</td>
<td>Action</td>
</tr>
<tr>
<td>----------------</td>
<td>--------</td>
</tr>
<tr>
<td>Inform #1: Strengthen and expand partnerships by broadening stakeholders’ awareness of NAEP and facilitating their use of NAEP resources</td>
<td></td>
</tr>
<tr>
<td>1. Develop and Sustain Partnerships // Identify What Partners Need to Expand Use and Utility of NAEP</td>
<td>Board staff</td>
</tr>
<tr>
<td></td>
<td>Board staff; Communications contractor</td>
</tr>
<tr>
<td></td>
<td>Communications contractor – Client Relationship Management tool (CRM)</td>
</tr>
<tr>
<td>2. Work with Partners to Increase Awareness and Use of NAEP</td>
<td>Board members; Board staff; NCES staff; Communications contractor</td>
</tr>
<tr>
<td>3. Focused Reporting of NAEP Results</td>
<td>Board staff; CRP contractor; Communications contractor</td>
</tr>
<tr>
<td>Responsibility</td>
<td>Action</td>
</tr>
<tr>
<td>----------------</td>
<td>--------</td>
</tr>
<tr>
<td></td>
<td>through partners</td>
</tr>
<tr>
<td>Board staff; Communications contractor</td>
<td>Produce quick graphics, videos, artifacts for dissemination</td>
</tr>
<tr>
<td>Highlight Contextual Data in Reporting</td>
<td>Board members; Board staff; NCES staff; Communications contractor; HumRRO technical support contract</td>
</tr>
<tr>
<td>Identify Opportunities to Promote Use of NAEP Data with Federal Datasets</td>
<td>Board members; Board staff; NCES staff</td>
</tr>
<tr>
<td>Inform #2: Increase opportunities to connect NAEP to administrative data and state, national, and international student assessments</td>
<td>Board members; Board staff; NCES staff</td>
</tr>
<tr>
<td></td>
<td>Responsibility</td>
</tr>
<tr>
<td>---</td>
<td>----------------</td>
</tr>
<tr>
<td>6.</td>
<td>Learn from Reporting of International Assessments <em>(Also, SV #8)</em></td>
</tr>
<tr>
<td></td>
<td>Board members; Board staff; NCES staff</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Inform #3: Expand the availability, utility, and use of NAEP resources, in part by creating new resources to inform education policy and practice</td>
</tr>
<tr>
<td>----</td>
<td>---------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>7.</td>
<td>Add Meaning to NAEP Achievement Levels</td>
</tr>
<tr>
<td>8.</td>
<td>Research Effective Uses of NAEP</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td>Develop New Tools for Audiences</td>
</tr>
<tr>
<td></td>
<td>Identify More User-Friendly Approaches to Presenting NAEP Results</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Board members; Board staff; NCES staff; Communications contractor</td>
<td>Construct custom portals for different subjects and/or types of users</td>
</tr>
</tbody>
</table>

|   | Create “Brief Case” Studies | Board staff; Communications contractor | Learn how NAEP used effectively by states and districts to serve as guide via compelling narratives in graphics, videos, two-pagers | Increased social media traffic; Number of “brief case studies” posted and re-posted | January 2018 - ongoing | Tennessee case study disseminated widely in June; Mississippi case study to be released soon |

| Board members; Board staff; NCES staff; Communications contractor | Create “menu of engagement” list of speakers, graphics, videos, artifacts that Board staff can offer partners | Artifacts developed for and posted by partners; Number of requests by partners; Number of activities | January 2018 - ongoing | Graphics and videos shared online and tagged to partners who retweet; |
**12. Facilitate Teacher Preparation Program Toolkit to Increase Access and Use of NAEP by Teachers**

| Board staff; Communications contractor | Meet with teacher educators to learn needs and interests | Develop tools and resources; Use of toolkits; User feedback | September 2018 | Met with AACTE Executive Director to initiate this idea |
| Communications contractor | Support development of toolkit by partners | Webpage on Governing Board website for teacher educators and preservice teachers | January 2019 |

**Inform #4:** Promote sustained dissemination and use of NAEP information beyond Report Card releases with consideration for multiple audiences and ever-changing multi-media technologies....

*Note: SV #4 permeates throughout the entire list of planned tasks and activities, so is not presented in separate rows.*

**Innovate #6:** Continue improving the content, analysis, and reporting of NAEP contextual data by considering the questions’ relevance, sensitivity, and potential to provide meaningful context and insights for policy and practice.

**13. Review Contextual Variables**

| Board members; Board staff | Review contextual variables to ensure relevance and importance | Greater use of contextual data; Updated variables | Ongoing | Reviewed core contextual items at May 2018 R&D meeting; Feedback registered and answered |
RELEASE PLAN FOR THE
NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS (NAEP)

The Nation’s Report Card: 2018 Technology and Engineering Literacy

The national results of the 2018 National Assessment of Educational Progress (NAEP) Technology and Engineering Literacy (TEL) Report Card will be released to the general public in Spring 2019 through a release event that will be based in the Research Triangle Park (RTP) area of North Carolina and webcast live for a national audience.

The RTP area (Raleigh-Durham-Chapel Hill) is a thriving hub of innovation in the area of STEM (science, technology, engineering, and mathematics). The event will be held at a location to fit the theme of the subject area and will utilize partners and resources in the region. The program will include a data presentation by the National Center for Education Statistics (NCES) on the national results at grade 8; moderation and comments by at least one Governing Board member; and participation from a diverse set of contributors, including experts in the technology and engineering fields, educators, and students, who can discuss the implications of the findings in the broader world of education and technology. To elevate the release event, the Board will pursue innovative and interactive ideas—which could include the use of video and other multimedia or technology—to demonstrate TEL tasks, to highlight the unique nature of the assessment, and to draw connections to STEM fields generally. This program, slated for 90 minutes, will include a conversational Q&A session with in-person attendees and the livestream audience.

The 2018 NAEP Technology and Engineering Literacy Nation’s Report Card presents real-world scenarios involving technology and engineering challenges. Students were asked to respond to questions aimed at assessing their knowledge and skill in understanding technological principles, solving technology and engineering-related problems, and using technology to communicate and collaborate. The report will include student and school survey responses about students’ experiences and their opportunities to learn in this subject area.

DATE AND LOCATION

The release events will occur in Spring 2019. The release date will be determined by the
Chair of the Reporting and Dissemination Committee, in accordance with Governing Board policy.

ACTIVITIES BEFORE THE RELEASE

In the months before the release event, the Governing Board will work with its communications contractor to identify a variety of potential technology-related partnerships in the higher-education, community, nonprofit, and private sectors to help promote the event, as well as assist with development of panels and other aspects of the program. Current and former Board members in North Carolina will also assist in terms of venue procurement, event promotion, and program development. National promotion, with a focus on social media efforts that both promote the event and inform the public about the TEL assessment, will be conducted as well.

In the days preceding the release, NCES will offer a conference call for appropriate media, and there will be an embargoed website with results available to Congressional staff, approved senior representatives of the National Governors Association and the Council of Chief State School Officers, and approved media. The goal of these activities is to provide stakeholders with a comprehensive overview of findings to help ensure accurate reporting to the public and deeper understanding of results.

REPORT RELEASE

The Commissioner of Education Statistics will publicly release the report at the NAEP website—http://nationsreportcard.gov—and at the scheduled time of the release event. An online copy of the report, along with data tools, questions, and other resources, will also be available at the time of release on the NAEP site. The Governing Board press release, the full and abridged versions of the 2018 NAEP Technology and Engineering Literacy Framework, and related materials will be posted on the Board’s web site at www.nagb.gov. The site will also feature links to social networking sites and multimedia material related to the event.

ACTIVITIES AFTER THE RELEASE

The Governing Board’s communications contractor will work with Board staff to coordinate additional post-release communications efforts—which could include a social media chat, major presentation, webinar, a video of the report’s contextual variables, infographics, or social media campaigns—to target communities and audiences with an interest in technology and engineering literacy. These efforts will involve identifying and working with stakeholders, and utilizing their channels and resources to reach their members, colleagues, and influencers. The goal of these activities is to extend the life of the results and provide value and relevance to stakeholders.
New Methods and Data Sources for Adding Context to NAEP Achievement Results

Enhancements to the NAEP questionnaires and greater access to data external to NAEP create opportunities for reporting of contextual variables. The 2017 mathematics and reading report cards featured expanded reporting on school and classroom contexts, as well as students’ attitudes toward school work. Recently released web reports that focused on the survey questionnaires have received substantial web traffic, even though the data are from 2015.

In this session NCES will discuss potential new data sources and approaches to expand on reporting of contextual variables in a cohesive and comprehensive manner. The R&D Committee will be asked for comment on the approaches presented.
Communication and Interpretation of Achievement Levels

At the November 2018 Governing Board meeting, COSDAM and the Reporting and Dissemination Committee will hold a joint meeting to discuss the two committees’ work on achievement levels for the National Assessment of Educational Progress (NAEP).

Background

From 2014 to 2016, the National Academies of Sciences, Engineering, and Medicine evaluated the NAEP achievement levels in mathematics and reading, which are the responsibility of the Governing Board. In their evaluation, the National Academies noted eight common uses of NAEP achievement levels, specifically:

- Trends or comparisons of successive cohorts, e.g., the percentage of students at or above Proficient in reading has increased over time;
- Comparison to a state assessment;
- Point-in-time comparisons across states, districts, or population groups, e.g., more students in state A who are at or above Proficient in reading compared to state B;
- Rank ordering states or districts;
- Comparison across population groups to examine performance gaps;
- Comparison across subject areas, e.g., more students perform at or above Proficient on mathematics than in reading;
- Comparison of before and after an action or policy implementation; and
- Relationships among achievement results and contextual data.

The evaluation recognized the usefulness and value of the achievement levels but made several important recommendations, most of which focus on the work of COSDAM as well as two that also address the work of the R&D Committee:

RECOMMENDATION 5: Research is needed to articulate the intended interpretations and uses of the achievement levels and collect validity evidence to support these interpretations and uses. In addition, research to identify the actual interpretations and uses commonly made by NAEP’s various audiences and evaluate the validity of each of them. This information should be communicated to users with clear guidance on substantiated and unsubstantiated interpretations.

RECOMMENDATION 6: Guidance is needed to help users determine inferences that are best made with achievement levels and those best made with scale score statistics. Such guidance should be incorporated in every report that includes achievement levels.
Since the release of these recommendations in November 2016, Governing Board staff and COSDAM members have started working to fulfill these recommendations. The draft revision of the Board policy on developing student achievement levels (scheduled for full Board action in November 2018) establishes an

“interpretative guide [which] shall accompany NAEP reports, including specific examples of appropriate and inappropriate interpretations and uses of the results” (Principle 3h).

This guide is intended for inclusion on the Nation’s Report Card website and on specific report card webpages. The guide will target stakeholders, such as media, policy advocates, members of the general public, educators, and policymakers. These groups may be familiar with both NAEP and achievement levels, but their understanding, interpretation, and use of achievement levels could be informed and improved with guidance from the Governing Board.

The Reporting and Dissemination Committee will collaborate with COSDAM on the development of this interpretative guide. The overarching question of the joint meeting will focus on the general approach the interpretative guide should take. This joint meeting also will elicit feedback on several specific features of the guide:

(1) the scope—what should be covered and what should not;
(2) the content—uses of achievement levels, value and usefulness of achievement levels;
(3) the language—non-technical, accessible; and
(4) the delivery—how the guide will be included with report cards.

If there is time, the conversation may extend to initial discussions of a statement on both the uses and usefulness of NAEP generally, not only of achievement levels specifically.

In addition, the committees should deliberate together on how to engage stakeholders on improving their use and interpretation of NAEP and achievement levels beyond the interpretative guide.
Closed Session: NAEP Assessment Schedule and Budget

Setting the NAEP Assessment Schedule is one of the Governing Board’s most important statutory responsibilities. Historically, the Governing Board has amended the NAEP Assessment Schedule to reflect legislative changes to NAEP’s authorization, new opportunities, and evolving expectations in what students should know and be able to do. According to the Governing Board’s General Policy on Conducting and Reporting NAEP, the Board “periodically establishes a dependable, publicly announced assessment schedule of at least ten years in scope. The schedule specifies the subject or topic (e.g., High School Transcript Study), grades, ages, assessment year, and sampling levels (i.e., national, state, TUDA) for each assessment.” The current Schedule of Assessments (attached) was approved in November 2015 and extends through 2024.

The Board’s Strategic Vision includes a priority to “Develop policy approaches to revise the NAEP assessment subjects and schedule based on the nation’s evolving needs, the Board’s priorities, and NAEP funding” (SV #9). To begin pursuing this strategic priority, Governing Board members engaged in small group and plenary discussions on this topic during several Board meetings over the past year. These discussions culminated in the adoption of a Resolution on Board Priorities for the NAEP Assessment Schedule (attached) at the March 2018 Board meeting.

During the May 2018 Board meeting, Governing Board members engaged in small group discussions to consider various approaches for implementing the assessment schedule priorities of utility, frequency, and efficiency. In its discussion, the Board generally agreed that it is desirable to increase state and TUDA administrations for Reading, Mathematics, Science, and Writing in particular.

During the August 2018 Board meeting, there was a plenary presentation and discussion of several potential approaches to increase the efficiency of the U.S. History, Civics, Geography, and Economics assessments. Relative benefits and costs of each approach were discussed, including potential implications for trends, achievement levels, and reporting. Several Board members noted that a consolidated social studies framework that would result in an overall social studies score and achievement levels may be too broad to be meaningful, helpful, or actionable. There was some support for separate assessments and trends in Civics and U.S. History to be maintained (even if the administrations are coordinated to produce results about interrelationships) and possibly prioritized over Geography and Economics.

Based on the Board’s previous discussions and operational considerations provided by NCES, Governing Board and NCES staff will share the rationale and projected costs associated with a proposed draft schedule through 2030.
National Assessment Governing Board Resolution on Priorities for the NAEP Assessment Schedule

Whereas, The Nation’s Report Card—also known as the National Assessment of Educational Progress (NAEP)—is mandated by Congress to conduct a national assessment and report data on student academic achievement and trends in public and private elementary schools and secondary schools (P.L. 107-279);

Whereas, the NAEP Authorization Act requires that NAEP be administered in public and private schools in reading and mathematics at least every 2 years in grades 4 and 8 and every 4 years in grade 12 and conduct the Long-Term Trend assessment in reading and mathematics for ages 9, 13, and 17;

Whereas, the NAEP Authorization Act specifies that beyond the requirements listed above, to the extent time and resources allow, NAEP shall assess and report achievement trends in additional subjects in grades 4, 8, and 12;

Whereas, the Every Student Succeeds Act mandates that states participate in the biennial reading and mathematics NAEP assessments in grades 4 and 8;

Whereas, Congress supported the establishment and expansion of the NAEP Trial Urban District Assessment (TUDA) to provide NAEP results for select large urban districts;

Whereas, NAEP provides national, state, and local policymakers and practitioners with consistent, external, independent measures of student achievement through which results across education systems can be compared at points in time and over time;

Whereas, the National Assessment Governing Board and the National Center of Education Statistics (NCES) continuously work to enhance NAEP’s form (e.g. transitioning to digital-based assessments) and content (e.g. the Technology and Engineering Literacy assessment) to reflect the modern expectations of what students know and can do;

Whereas, Congress authorized the National Assessment Governing Board to determine the NAEP subjects to be assessed;

Whereas, it is the National Assessment Governing Board’s policy, in consultation with NCES, to periodically establish a dependable, publicly announced NAEP Schedule of Assessments spanning at least ten years, and specifying the subjects, grades, ages, assessment years, sampling levels (e.g., national, state, TUDA), and introduction of new and revised frameworks for each assessment;

Whereas, on November 18, 2016 the National Assessment Governing Board unanimously adopted its Strategic Vision which included a priority to “Develop policy approaches to revise the NAEP assessment subjects and schedule based on the nation’s evolving needs, the Board priorities, and NAEP funding”;
Therefore, as the National Assessment Governing Board anticipates extending the NAEP Schedule of Assessments into the future, it will uphold all of the aforementioned requirements and make decisions informed by each of the following priorities to ensure NAEP results are impactful and policy-relevant:

- **Utility** – include more voluntary state and Trial Urban District Assessments and continue to align the schedule of NAEP administrations with international assessments in the same subjects to enable actionable comparisons of districts, states, and other nations;

- **Frequency** – commit to assess subjects other than reading and mathematics at least every 4 years to provide additional measures of student academic progress at regular intervals; and

- **Efficiency** – find cost-effective ways to administer NAEP while to the degree possible maintaining a breadth of subjects on the schedule in order to continue reporting progress in student achievement;

Furthermore, the National Assessment Governing Board recognizes that any change to the NAEP Schedule of Assessments requires consideration of the fiscal, technical, and operational implications.
Reading
- NAEP legislation specifies every 2 years at grades 4 and 8 for nation and states; NCLB/ESSA requires states to partake
- NAEP legislation specifies every 4 years at grade 12 for nation
- Administration has included voluntary TUDAs for grades 4 and 8 since 2002
- Administered at national level only for grade 12, and for 11-13 states voluntarily participated in 2009 and 2013
- Grade 12 assessment used to estimate % of students academically prepared for college
- Current trend lines begin in 1992
- Administration coincides with PIRLS (grade 4) once every 10 years

Math
- NAEP legislation specifies every 2 years at grades 4 and 8 for nation and states; NCLB/ESSA requires states to partake
- NAEP legislation specifies every 4 years at grade 12 for nation
- Administration has included voluntary TUDAs for grades 4 and 8 since 2003
- Administered at national level only for grade 12, and for 11-13 states voluntarily participated in 2009 and 2013
- Grade 12 assessment used to estimate % of students academically prepared for college
- Current trend lines begin in 1990 for grades 4 and 8; 2005 for grade 12
- Administration coincides with every administration of TIMSS (4 year cycle)

Science
- Has been administered approximately every 4 years at all 3 grades
- Administered to the nation, states, and (usually) voluntary TUDAs for grades 4 and 8
- Administered at national level only for grade 12
- Current trend lines begin in 2009
- Since 2011, administration has coincided with every administration of TIMSS

Writing
- Has been administered approximately every 4 years at grades 8 and 12; much less frequently at grade 4
- Under current framework (beginning with 2011 administration), has been administered to the nation only
- Previous framework included administration to states and voluntary TUDAs in 1998 (states only), 2002, 2007
History
- Has been administered at the national level approximately every 4 years at grade 8; less frequently at grades 4 and 12

Civics
- Has been administered at the national level approximately every 4 years at grade 8; less frequently at grades 4 and 12

Geography
- Has been administered at the national level approximately every 4 years at grade 8; less frequently at grades 4 and 12

Technology and Engineering Literacy (TEL)
- Has been administered at national level for grade 8 only in 2014 and 2018
- Framework covers all 3 grades

Economics
- Framework covers grade 12 only
- Has been administered at national level in 2006 and 2012

Arts
- Framework covers all 3 grades but administered at national level for grade 8 only
- Framework includes 4 areas (Dance, Music, Visual Arts, and Theatre) but only Music and Visual Arts have been included in operational assessment
- New framework is needed for transition to DBA; not feasible to complete in time for 2024 administration

Foreign Language
- Framework to measure Spanish language proficiency adopted in 2000
- Pilot test conducted in 2003 but assessment never administered operationally

High School Transcript Study
- Supplemental data collection to grade 12 Math and Science administrations
- NCES has been working to determine the feasibility of conducting this study for grade 8 and at the state level

Long-Term Trend (LTT)
- Legislation notes continuing for Reading and Math, but no periodicity specified
- Periodicity has varied but generally has been at least every 4 years until 2012
The National Assessment of Educational Progress (NAEP) Authorization Act established the National Assessment Governing Board to set policy for NAEP, including determining the schedule of assessments. (P.L. 107-279)

<table>
<thead>
<tr>
<th>Year</th>
<th>Subject</th>
<th>National Grades Assessed</th>
<th>State Grades Assessed</th>
<th>TUDA Grades Assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>U.S. History*</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Civics*</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Geography*</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TECHNOLOGY AND ENGINEERING LITERACY</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>Reading*</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td></td>
<td>Mathematics*</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td></td>
<td>Science**</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td>2016</td>
<td>Arts*</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td>Reading</td>
<td>4, 8</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td></td>
<td>Mathematics</td>
<td>4, 8</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td></td>
<td>Writing</td>
<td>4, 8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>U.S. History</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Civics</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Geography</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Technology and Engineering Literacy</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2019</td>
<td>Reading</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td></td>
<td>Mathematics</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td></td>
<td>Science</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td></td>
<td>High School Transcript Study</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2020</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2021</td>
<td>Reading</td>
<td>4, 8</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td></td>
<td>Mathematics</td>
<td>4, 8</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td></td>
<td>Writing</td>
<td>4, 8, 12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2022</td>
<td>U.S. HISTORY</td>
<td>8, 12</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CIVICS</td>
<td>8, 12</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>GEOGRAPHY</td>
<td>8, 12</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Economics</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Technology and Engineering Literacy</td>
<td>8, 12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2023</td>
<td>Reading</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td></td>
<td>Mathematics</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td></td>
<td>Science</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td></td>
<td>High School Transcript Study</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2024</td>
<td>ARTS</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FOREIGN LANGUAGE</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Long-term Trend</td>
<td>~</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**NOTES:**
* Assessments not administered by computer. Beginning in 2017 all operational assessments will be digitally based.
** Science in 2015 consisted of paper-and-pencil and digital-based components.
~Long-term Trend (LTT) assessments sample students at ages 9, 13, and 17 and are conducted in reading and mathematics.
Subjects in **BOLD ALL CAPS** indicate the year in which a new framework is implemented or assessment year for which the Governing Board will decide whether a new or updated framework is needed.
One of the Governing Board’s most important legislative responsibilities is developing the NAEP achievement levels. The history of NAEP achievement level setting (also known as standard setting) over the past 30 years is complex.

When the Governing Board developed its initial policy for developing NAEP achievement levels in 1990, the field of standard setting was much less developed than it is today and most experiences were in the area of professional licensure. At that time, very little standard setting had been conducted in K-12 education, particularly at the national level. The Governing Board’s initial efforts to develop achievement levels for NAEP were thus groundbreaking, and the Board’s research and practices in this area over the past three decades have had a large impact on how standard setting is conducted today in K-12 education.

Early evaluations of the NAEP achievement levels (conducted in the 1990s) were critical of the Governing Board’s procedures and of judgmental standard setting in general. When the initial NAEP legislation was re-authorized in 1994, it included language that the NAEP achievement levels should be “used on a developmental basis until the [NCES] Commissioner determines, as a result of an evaluation... that such levels are reasonable, valid, and informative to the public” (Public Law 103-382). The next re-authorizations substituted the term “trial” for “developmental” (Public Laws 107-110 and 107-279).

In 1999, a report of an evaluation conducted by the National Academy of Sciences stated that the process for setting NAEP achievement levels was “fundamentally flawed.” This conclusion was disputed by many technical experts, and several prominent researchers issued a response criticizing the evaluation. Although the Governing Board has adopted new approaches to developing the NAEP achievement levels since the 1990s, the original methodology (called “modified Angoff”) is still considered a reputable approach to standard setting and remains in wide use today. For example, a version of the modified Angoff methodology was used to set the achievement levels for the Partnership for Assessment of Readiness for College and Careers (PARCC).

The more recent evaluations of the NAEP achievement levels have recognized their value and have been more positive about the standard setting methodology used by NAEP. The 2009 evaluation of the NAEP program as a whole included a focus on the Mapmark achievement-level setting activities for the 2005 grade 12 mathematics assessment and noted:
“Based on our analysis of the procedural, internal, and external evidence pertaining to the validity of the process of setting achievement level standards on the 2005 Grade 12 NAEP Mathematics assessment, we conclude that the procedure was sound, followed recommendations for best practices in the area of standard setting, and involved multiple quality control checks to support the defensibility of the process. The validity of any type of test score interpretation is not something that can be unequivocally established. However, the multiple sources of validity evidence we analyzed for this exam lead us to conclude the standards set on the 2005 Grade 12 NAEP Mathematics assessment are valid for the purposes of describing the performance of 2005 Grade 12 students with respect to the NAEP achievement level descriptors” (pages 2-58 to 2-59).

The report included several recommendations to further improve the process.

The 2016 evaluation of NAEP achievement levels, conducted by the National Academies of Sciences, Engineering, and Medicine, focused on the NAEP mathematics and reading achievement levels for grades 4, 8, and 12. This evaluation report, presented to the Governing Board at its November 2016 meeting, stated, “During their 24 years [the achievement levels] have acquired meaning for NAEP’s various audiences and stakeholders; they serve as stable benchmarks for monitoring achievement trends, and they are widely used to inform public discourse and policy decisions. Users regard them as a regular, permanent feature of the NAEP reports” (page Sum-8). This evaluation included several recommendations, and the Board issued a formal response noting its planned actions in December 2016.

One important aspect of the Board’s response to the 2016 evaluation was a commitment to update the guidance provided in the Board policy statement on Developing Student Achievement Levels for NAEP. The Committee on Standards, Design and Methodology (COSDAM) has been working to update this policy guidance over the past 1.5 years, and the Board recently sought public comment on the proposed policy statement. The strong response to the public comment comprised both strong support for the Governing Board’s process and policies, and concerns about potential misunderstandings of NAEP achievement level labels.

We have assembled a panel to provide some history and perspectives on the complex history of the NAEP achievement levels. Moderated by Governing Board member and standard setting expert Dr. Gregory Cizek, panelists will provide brief remarks and will engage in a question and answer session with the Board. The attached materials include panelist biographies, and a history of NAEP achievement level setting written by the Governing Board’s first Assistant Director for Psychometrics, Mary Lyn Bourque, as part of the Governing Board’s 20th anniversary event in 2009 (prior to the most recent evaluations).
Dr. David Driscoll served as the 22nd Commissioner of Elementary and Secondary Education in the Commonwealth of Massachusetts from 1998 to 2007. He also served as president of the Council of Chief State School Officers. Dr. Driscoll was a member of the National Assessment Governing Board from 2006 to 2014 and served as Chair from 2009 to 2014.

Dr. Driscoll’s career in public education and educational leadership spans more than four decades. A former secondary school mathematics teacher, he was named Melrose Assistant Superintendent in 1972 and Superintendent of Schools in the same community in 1984. He served in that role until 1993, when he was appointed Massachusetts Deputy Commissioner of Education, just days after the state’s Education Reform Act was signed into law. He became Interim Commissioner of Education in July 1998 and was named Commissioner in March 1999.

As Deputy Commissioner, Dr. Driscoll played several key leadership roles, including as the Principal Investigator for the National Science Foundation’s Partnerships Advancing the Learning of Mathematics and Science (PALMS) in Massachusetts. As Commissioner, Dr. Driscoll oversaw the development of the state’s curriculum frameworks, implementation and expansion of the Massachusetts Comprehensive Assessment System, development of the Massachusetts School and District Accountability System, and development and administration of the Educator Certification Test and new licensure regulations. These initiatives and others led to consistent annual improvement in student achievement as measured by state tests, national measures, and international tests. In 2005, Massachusetts was the first state to earn the highest scale score in the nation on all four National Assessment of Educational Progress exams.

Dr. Driscoll earned his bachelor’s degree in mathematics at Boston College, master’s degree in educational administration from Salem State College, and his Ph.D. in education administration from Boston College.

Chester E. Finn Jr. has devoted his career to improving education in the United States. As a senior fellow at Stanford’s Hoover Institution, former chairman of Hoover’s Task Force on K-12 Education, member of the Maryland State Board of Education and of Maryland’s Commission on Innovation and Excellence in Education, and Distinguished Senior Fellow & President Emeritus of...
the Thomas B. Fordham Institute, his primary focus is reforming primary and secondary schooling.

Finn served as Fordham’s President from 1997 to 2014, after many earlier roles in education, academe, and government, including professor of education and public policy at Vanderbilt University, US assistant secretary of education, and legislative director for Senator Daniel Patrick Moynihan.

A native of Ohio, he holds an undergraduate degree in US history, a master’s degree in social studies teaching, and a doctorate in education policy, all from Harvard University.

Finn has served on numerous boards, currently including the National Council on Teacher Quality and the Core Knowledge Foundation. From 1988 to 1996, he served on the National Assessment Governing Board, including two years as its chair.

Author of more than four hundred articles and twenty books, Finn’s latest (coauthored with Jessica Hockett) is Exam Schools: Inside America’s Most Selective Public High Schools.

Laura S. Hamilton is a senior behavioral scientist and distinguished chair in Learning and Assessment at the RAND Corporation. She also serves as faculty member at the Pardee RAND Graduate School and adjunct faculty member in the University of Pittsburgh’s Learning Sciences and Policy program. Her research addresses topics related to social and emotional learning, educational assessment, accountability, school leadership, the implementation of curriculum and instructional reforms, and education technology. Recent projects include a study of a social and emotional learning intervention for elementary schools and afterschool programs, a review of evidence of social and emotional learning interventions, the development of a database of measures of student interpersonal and intrapersonal competencies, and a review of indicators of standards-aligned science and mathematics instruction. She also co-directs the American Educator Panels, RAND’s nationally representative survey panels of teachers and principals. She served as a member of the committee that revised the Standards for Educational and Psychological Testing and is currently a member of the steering committee for the Center for Academic, Social, and Emotional Learning’s Assessment Work Group. She has also served on several state and national panels on topics related to assessment, accountability, and educator evaluation including the National Academies of Sciences, Engineering, and Medicine Committee on the Evaluation of NAEP Achievement Levels and its Committee on Developing Indicators of Education Equity. Hamilton earned a Ph.D. in educational psychology and an M.S. in statistics from Stanford University.
Marc S. Tucker is the President and CEO of the National Center on Education and the Economy, and has been a leader of the standards-driven education reform movement for many years. Mr. Tucker created New Standards, a 23-state consortium designed to develop internationally benchmarked student performance standards and matching student examinations. He authored the 1986 Carnegie Report, A Nation Prepared: Teachers for the 21st Century, which called for a restructuring of America’s schools based on standards; created the National Board for Professional Teaching Standards; created the Commission on the Skills of the American Workforce and co-authored its report, America’s Choice: high skills or low wages!, which called for a new high school leaving a certificate based on standards; and, was instrumental in creating the National Skill Standards Board and served as the chairman of its committee on standards and assessment policy.

With Ray Marshall, Mr. Tucker co-authored Thinking for a Living: Education and the Wealth of Nations, selected by Business Week as one of the ten best business books of 1992; with Judy Codding, co-authored Standards for Our Schools: How to Set Them, Measure Them, and Reach Them, published in 1998; and co-edited The Principal Challenge, 2002. Mr. Tucker created the National Institute of School Leadership, a state-of-the-art executive development program for school leaders. Mr. Tucker was the lead author of Tough Choices or Tough Times, the report of the New Commission on the Skills of the American Workforce. In 2014, the Education Commission of the States awarded Mr. Tucker the James Bryant Conant award for his outstanding individual contribution to American education. Mr. Tucker currently has an appointment as a Visiting Distinguished Fellow at the Harvard Graduate School of Education.

Mary Lyn Bourque

March 2009

Paper Commissioned for the 20th Anniversary of the National Assessment Governing Board 1988–2008

Mary Lyn Bourque is President of Mid Atlantic Psychometric Services of Leesburg, VA. From 1989 to 2000 she was Assistant Director for Psychometrics of the National Assessment Governing Board.
Introduction

A history of the achievement levels for the National Assessment of Educational Progress (NAEP) could be documented in different ways. A chronological history might be the obvious approach, but in the author’s view, that could miss the most salient aspects of setting student performance standards on an assessment like NAEP. Further, new initiatives are fostered within a social and political context that is critical to understanding the initiative’s development, direction, and destiny. Therefore, rather than following a strict chronology, the author has adopted an issues-based approach, describing first the contextual issues that surrounded NAEP during the beginnings of achievement levels in the late 1980s and early 1990s. The paper will then discuss the critical initial implementation decisions made by the National Assessment Governing Board as it developed the student performance standards for NAEP. Then it will follow new implementation issues into the first decade of the 21st century; finally, it will discuss the broad impact of the achievement levels on American education.

Contextual Issues: Early NAEP

NAEP emerged as a federally supported program in the late 1960s, often pinpointed at October 1969 when the U.S. Office of Education (USOE) awarded a 1-year, $2 million grant to the Education Commission of the States (ECS) to support the planning and initial assessments of NAEP. Prior to that time (during the initial planning phases and dating from spring 1963), support for NAEP came from private sources, including the Carnegie and Ford Foundations, with encouragement by the USOE. Two years later (August 1971), however, the USOE transferred administrative responsibility to the National Center for Education Statistics (NCES), where it remains today (though with some legislative modifications). It took another two years (September 1973) for the USOE to announce that NAEP’s funding arrangement would move from a grant to a contract. The fiscal year 1975 budget request by President Nixon included $7 million for NAEP. Eventually, Congress appropriated $3 million.

Interestingly, NAEP emerged as an initiative that Francis Keppel, then U.S. Commissioner of Education, believed would contribute to fulfilling the intent of the legislation that created the USOE 100 years earlier. That legislation required that the Commissioner of Education provide an annual report to Congress on the state of American education. For the first 100 years the annual reports focused on student-teacher ratios, per capita spending on education, number of classrooms and teachers, and teacher salaries, but provided little or no information on the outcomes of education; that is, what students know and can do. Keppel thought that a national assessment of students conducted on a regular basis could fill that void.

The initial design of the NAEP program was developed carefully to protect the rights of the states. It virtually precluded using NAEP as a lever for policy changes in American education: NAEP sampled students by age, not by grade; NAEP reported results at the regional level of aggregation, not by states or districts; and NAEP participation was strictly voluntary. A testimony to the wisdom of the original group of crafters, and to the adage that “change comes slowly,” NAEP is still making its presence felt as a partner to the change embodied in the No Child Left Behind (NCLB) legislation (P.L. 107–110).
The early years of NAEP planning occurred during the post-Sputnik era. The U.S.S.R. launched its first satellite orbiting the earth in 1957. This event sent shockwaves throughout the western world, particularly the United States, which viewed itself as a world leader. So while there was an urgent need to know what American students know and can do, there was an equally urgent need to critically evaluate the American system of education and its relationship to our advancement as a nation. This was the era of the National Science Foundation’s support for science teachers and foreign language learning as well. In short, it was a time of taking stock in how we fared as a nation, what future generations of American citizens could and would accomplish, and how we could improve education by making it more accountable at all levels.

Internationally, we were (in the 1960s) and still are (in 2009) one of the only developed countries that does not have a centralized structure of elementary and secondary education. Contrary to other institutions, education in the United States has never been “nationalized.” Two hundred years after the founding fathers, the United States has no national curriculum, no national content or performance standards, no national testing of individual students, no national accountability, no national expectations for a high school diploma (McNeil, 2008), and no national training program or entry standards for teachers or administrators. In addition to more than 50 publicly funded state education systems and nearly 15,000 local education agencies, there are numerous private schools, charter schools, alternative schools, home schools, and other forms of schooling that satisfy compulsory education laws enacted by states, not by the federal government. The founding fathers awarded full responsibility for the education of the nation’s children to the states, and federal intrusion into such matters has always been viewed with suspicion.

The mix of the domestic social and historical contexts of the 1960s and 1970s resulted in a National Assessment that developed slowly. From funding issues to administrative issues, NAEP struggled to develop into a first-class assessment program without a broad-based federal infrastructure to support it. Consensus came at a snail’s pace from the various public groups, especially Congress. However, by the mid-1980s the political climate began to change. A Nation at Risk (National Commission on Excellence, 1983) was published, the Educational Testing Service (ETS) was awarded the new NAEP contract, and a New Design for a New Era (Messick, Beaton, and Lord, 1983) provided a whole new framework for NAEP.

During the first 20 years of NAEP, much was accomplished due to the diligence of its contractors and those at the federal level responsible for its administration. Assessments were developed and administered to samples of K–12 students in a variety of subjects, including science, citizenship, reading, writing, and computer competency; in addition, out-of-school samples and a young adult literacy study (16- to 25-year olds) were implemented. However, despite some progress, the direction of NAEP was too dispersed among contractors, federal agencies, and interested public groups to allow the program to reach its full potential as the nation’s leading indicator of academic progress. Someone needed to be “in charge” if NAEP was to move forward.
Contextual Issues: Early Governing Board

By the mid-1980s, states began to realize that better reporting mechanisms were needed to measure student progress. Comparisons of SAT and ACT college admission scores state by state were no longer satisfactory. The National Governors’ Association (Alexander, 1986) called for better “report cards” that could more accurately compare states’ performances to each other and to the nation.

The Alexander-James Panel, appointed by William Bennett (then Secretary of Education), was a response to the governors’ call. The 22-member panel report issued in 1987 (Alexander and James, 1987), along with commentary by the National Academy of Education (Glaser, 1987), ultimately formed the basis of the reauthorization of the Elementary and Secondary Education Act, also known as the Hawkins-Stafford Elementary and Secondary School Improvement Amendments of 1988. The Governing Board was created by this federal legislation (P.L. 100–297, 1988), which was signed by President Reagan in April 1988.

The new legislation covering NAEP was the first attempt to put someone “in charge,” at least in the policy sense. In theory, the law gave the Governing Board responsibility for program policy, NCES would oversee program administration, and the NAEP contractors would provide technical expertise. The law set NAEP on what some refer to as a three-legged stool for support: a tripartite, shared responsibility where the distinction between policy and administration was not clear, and policy and technical issues overlapped considerably. In theory it sounded logical. But in fact, some of these distinctions are still being worked out today, some 20 years later.

The Governing Board took its legislative charge seriously. Clearly, its most controversial responsibility was that of setting student performance standards for NAEP, which the Board decided to term achievement levels. P.L. 100–297 makes more than one reference to what could be interpreted as student performance standards. One part of the statute directs the Board to “identify appropriate achievement goals for each age and grade in each subject area to be tested (Sec. 3403, (6)(A).” In another section the law states that “Each learning area assessment shall have goal statements devised through a national consensus approach (Sec. 3403, (6)(E).” Both directives could be interpreted to mean different things and, further, both could be interpreted differently by different groups. In order to better fulfill the intent of the law, the Board leaned heavily on the National Academy of Education (NAE) commentary on the Alexander-James report, which was the basis of the legislation (Alexander and James, 1987; Glaser, 1987). In that document the Academy panel argued that “for each content area NAEP should articulate clear descriptions of performance levels, descriptions that might be analogous to such craft rankings as novice, journeyman, highly competent, and expert. Descriptions of this kind would be extremely useful to educators, parents, legislators, and an informed public (Glaser, 1987).”

During the first 18 months after the Board’s formation, it developed a policy statement on NAEP student performance standards. The Board agreed to adopt three achievement levels (Basic, Proficient, and Advanced) for each grade and subject area assessed by NAEP. In arriving at that decision, the Board researched different options, including whether the legislative phrase “identifying appropriate achievement goals [emphasis added]” was synonymous with setting “targets” on NAEP, or whether the law’s intent was to identify the content that groups of
students should know and be able to do if they reached the standard. The Board ultimately rejected the notion of “targets,” believing that setting passing rates for the decentralized education system of the United States was beyond the policy competence of a national Board.

The Legislative History of Achievement Levels

Before launching into an issues-based history of the NAEP achievement levels, it might be beneficial to briefly summarize the legislative history of the achievement levels (figure 1). Before 1988, the NAEP legislation was mute on such things as performance standards. This early legislation was content to cover such things as which subject areas would be assessed, which samples would be assessed, where the resources would come from, etc.

The first NAEP legislation that spoke to performance standards was P.L. 100–297 of 1988, which established the Governing Board to set policy and required it to identify “appropriate achievement goals for each…grade… (and) subject area” to be assessed. Although the law was ambiguous about exact expectations, the Board understood that it was charged with developing student performance standards. The legislation did not spell out how, but left that to the wisdom of the broadly representative Board.

The 1994 NAEP legislation, Improving America’s Schools Act of 1994 (P.L. 103–382), came out of the same congressional session as the Goals 2000 law, Educate America Act (P.L. 103–227). The latter codified the education goals promulgated by the Charlottesville Summit and reflected the early policy definitions adopted by the Board, which, in turn, borrowed from the language used at Charlottesville. The National Education Goal 3 stated in part that, “All students will leave grades 4, 8, and 12 having demonstrated competency over challenging subject matter… so they may be prepared for responsible citizenship, further learning, and productive employment in our nation’s modern economy.”

In the 1994 NAEP legislation, the achievement levels were judged “developmental” for the first time. This came as a result of two outside evaluations, which concluded that the student performance standards on NAEP did not meet technical expectations and should not be used as the primary means for reporting NAEP. (National Academy of Education, 1993; U.S. General Accounting Office, 1993).

The current legislation (P.L. 107–110), adopted in 2001, places the achievement levels within the accountability framework of NCLB, though they still are to be used by NAEP on a “trial basis.” States receiving federal education aid must participate in the National Assessment. They must also set standards on their own state assessments using the terms of the NAEP achievement levels (Basic, Proficient, and Advanced), though there is no requirement to have the same academic content or rigor.
Figure 1
Legislative History of Achievement Levels: 1988–2001

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(6)(A)(ii) “identifying appropriate achievement goals for each age and grade in each subject tested under the National Assessment;”</td>
<td>Sec. 411 (e) STUDENT PERFORMANCE LEVELS “PERFORMANCE LEVELS. The National assessment Governing Board, established under section 412, shall develop appropriate student performance levels for each age and grade in each subject area to be tested under the National Assessment. DEVELOPMENT OF LEVELS. Such levels shall be – Devised through a national consensus approach, providing for active participation of teachers, curriculum specialists, local school administrators, parents, and concerned members of the general public; Used on a developmental basis until the Commissioner determines, as a result of an evaluation…, that such levels are reasonable, valid, and informative to the public; Updated as appropriate. In using such levels on a developmental basis, the Commissioner and the Board shall ensure that reports that use such levels do so in a manner that makes clear the developmental status of such levels. REPORTING – After determining that such levels are reasonable, valid, and informative to the public, as a result of an evaluation …, the Commissioner shall use such levels or other methods or indicators for reporting results of the National Assessment and State Assessments.”</td>
<td>Sec. 602 (e) STUDENT ACHIEVEMENT LEVELS “ACHIEVEMENT LEVELS. The National Assessment Governing Board shall develop appropriate student achievement levels for each age and grade in each subject area to be tested under assessments authorized under this section, except the trend assessment… DETERMINATION OF LEVELS – In general such levels shall be determined by identifying the knowledge that can be measured and verified objectively using widely accepted professional assessment standards; and developing achievement levels that are consistent with relevant widely accepted professional assessment standards and based on the appropriate level of subject matter knowledge for grade levels to be assessed, or the age of the students, as the case may be. NATIONAL CONSENSUS APPROACH – After the determinations described in subparagraph (A), devising a national consensus approach. TRIAL BASIS – The achievement levels shall be used on a trial basis until the Commissioner determines, as a result of an evaluation…, that such levels are reasonable, valid, and informative to the public. STATUS – The Commissioner and the Board shall ensure that reports using such levels on a trial basis do so in a manner that makes clear the status of such levels. UPDATES – Such levels shall be updated as appropriate by the National Assessment Governing Board in consultation with the Commissioner. REPORTING – After determining that such levels are reasonable, valid, and informative to the public, as a result of an evaluation …, the Commissioner shall use such levels or other methods or indicators for reporting results of the National assessment and State assessments. REVIEW – The National Assessment Governing Board shall provide for a review of any trial student achievement levels under development by representatives of State educational agencies or chief State school officers…”</td>
</tr>
</tbody>
</table>
The differences between these pieces of authorizing legislation are quite nuanced in one sense. At first blush, they all look pretty much alike, although the most current one is far more explicit than the initial one. On the other hand, there are distinctions that affect how the achievement levels are viewed, used, and maintain their currency. As the paper discusses the various issues below, these differences will be highlighted.

**Implementation Decisions: The Beginning of Achievement Levels**

The Board faced many implementation questions during the early years of developing achievement levels, including the number of levels and what they should be called, a description of the achievement levels, the methodology to be used to develop them, the composition of standard-setting panels, how to report student performance standards, and the transition from anchor levels to achievement levels. The following sections will examine each of these topics.

**How Many Levels?**

This question was one of the first to come up in developing Board policy. The 1970s and 1980s experienced the growth of the minimal competency movement. The historical result of that movement was one of mediocrity. Indeed, standards were set, usually a single passing score and usually on tests of basic skills, and at a level to ensure most students would “pass.” However, the minimum competency paradigm would not work for NAEP in the context of *A Nation at Risk*, the Charlottesville summit, and the push for international competitiveness in a global economy. On the other hand, setting standards too high on a challenging test such as NAEP could result in irrelevance for NAEP and, more importantly, in a lack of involvement in the completely voluntary National Assessment. Furthermore, it would be far more beneficial to the states (at whose behest improved reporting of academic student performance was initiated) to be able to describe the performance of students across the whole distribution, not just at a single “passing score.” States wanted to know how all their students performed—those at the top of the distribution as well as those in the middle and lower ranges. Indeed, they wanted to know if some of their students met or even exceeded expectations, and if some of their students did not meet any standards or were in the “almost there” category. This called for more than just a “pass.”

Although the Board considered reporting a single level of performance, it was convinced to consider more than one level, especially by the late Albert Shanker, then President of the American Federation of Teachers. In fact, NAEP was ideally suited to employing multiple levels since the item pool used at any one grade level was substantial (e.g., the 1990 math assessment used nearly 200 items to measure grade 8 performance), and the use of large item pools was a necessary condition for having multiple standards. In the final analysis, the Board adopted three levels—Basic, Proficient, and Advanced—that were defined by Governing Board policy and applicable to all subject areas and grades.

Why three? The reasons are both practical and technical. The technical reason is that the scale range would probably not support more than nine levels of performance (three grades x three levels) on a single cross-grade scale and still provide clear distinctions between them, as Board policy called for. Although this was an empirical question at the time, NAEP data have
borne this out. The practical reason is somewhat related. Immediately before the adoption of the Board’s policy, NAEP had been using the ETS-developed NAEP “anchor levels”—one in the middle of the distribution, two higher, and one lower. Since one of the purposes of adopting the achievement levels was to “improve the form and use of NAEP,” three levels at each grade seemed to be the “Goldilocks approach” that would be “just right.”

Other labels were considered during the policy development process, including the suggestions found in the NAE commentary (Glaser, 1987) on the Alexander-James report (1987). These included such labels as “novice, journeyman, highly competent, and expert.” There was also much discussion about using the label “proficient” precisely because the anchor points were collectively referred to as NAEP proficiency. Similarly, the Board entertained a numerical labeling (e.g., Levels 1, 2, 3, etc.) and the terms “fundamental” and “master” for the Basic and Proficient levels, respectively. However, none of these seemed to meet the legislative charge of “improving the form and use of NAEP,” nor did they comport with the descriptions of the standards the Board had in mind. The Board wanted a more descriptive label that corresponded to the content of the levels of expectations envisioned by the policy. In the final analysis, alternatives were eventually set aside in favor of the current identifiers.

Policy Definitions

Who should set the standards on NAEP does not seem to be a salient question today, but in 1989 it truly was an issue. The Board struggled in the early days to find a way to develop a process for standard setting that invoked the Board’s role as the policy body for NAEP. The Board believed that, since it was the legal entity responsible for setting standards under the federal statute, it should set the expectations for “how good is good enough.” By the 1992 mathematics standard-setting initiative the Board had arrived at a three-prong solution: (1) the Board would develop policy definitions (PDs) that articulate the expectations for each level in general terms, (2) the PDs would be operationalized into grade- and subject-specific statements of performance levels for use in the standard-setting process (achievement level descriptions, or ALDs), and (3) the Board would be the final arbiter of the recommendations provided on all aspects of the levels, including the ALDs, the cut scores, and the exemplar items.

The first general definitions of performance (i.e., PDs) were developed to ensure that the standards from subject area to subject area, and from grades 4 to 8 to 12, were aligned. In other words, the Board did not want different standards (easier or harder) in reading than in math, or in science than in writing, nor did they want different standards in grades 4, 8, or 12. The first definitions were called policy definitions to distinguish them from the achievement level descriptions developed later. The original policy definitions appear in figure 2.
**Original Policy Definitions**

*(Excerpt from the Governing Board Policy (May 10, 1990))*

**Proficient.** This central level represents solid academic performance for each grade tested—4, 8, and 12. It will reflect a consensus that students reaching this level have demonstrated competency over challenging subject matter and are well prepared for the next level of schooling. For grade 12, the proficient level will encompass a body of subject-matter knowledge and analytical skills, cultural literacy, and insight that all high school graduates should have for democratic citizenship, responsible adulthood, and productive work.

**Advanced.** This higher level signifies superior performance beyond proficient grade-level mastery at grades 4, 8, and 12. For grade 12, the advanced level will show readiness for rigorous college courses, advanced technical training, or employment requiring advanced academic achievement. As data become available, it may be based in part on international comparisons of academic achievement and may also be related to Advanced Placement and other college placement tests.

**Basic.** This level, below Proficient, denotes partial mastery of knowledge and skills that are fundamental for proficient work at each grade tested. For grade 12, this will be higher than minimum competency skills (which normally are taught in elementary and junior high schools) and will cover significant elements of standard high school level work.
The opening page of the Board’s policy statement, unanimously approved on May 11, 1990, stated:

“The National Assessment Governing Board is not authorized to establish any overarching national goals for education. It does have authority to define levels of achievement that will serve as ‘appropriate achievement goals’ on National Assessment exams . . . Hence the achievement levels defined by the Board will be used for reporting group data and making it more meaningful (National Assessment Governing Board, 1990, p. 1).”

That initial policy statement goes on to say:

“The proposed achievement levels will add to assessment frameworks and objectives the specific definitions of basic, proficient, and advanced achievement at each grade tested, which are based on the content of the National Assessment exams. These are not broad general goals of education or curriculum, but substantive descriptions of levels of achievement tied firmly to National Assessment questions and objectives (National Assessment Governing Board, 1990, p. 6).”

These content-specific statements became known as the achievement level descriptions. The title of the early Board policy, “Setting Appropriate Achievement Levels for the National Assessment of Educational Progress” indicates that the Board initiative was never intended to reflect national goals for school subjects. Level setting was conceptualized as follows:

- limited to better reporting the results of the NAEP survey
- limited to three grade levels (4, 8, and 12)
- limited to specific subjects areas selected by the Board
- limited in scope, reflecting a limited assessment framework

The levels would answer the question, “How good is good enough, on NAEP?”

Despite policy arguments regarding the limited nature of the NAEP standards, the original policy definitions were ambitious. This was understandable given the educational climate existing at the time. Congress passed the Hawkins-Stafford Act because better information was needed on the performance of American students. States wanted better reporting as well, and it was thought that making the NAEP scale understandable to the public would be a way of reaching that goal. Grade 12 was singled out in all three original definitions because grade 12 is generally viewed as the gatekeeper for any postsecondary choices, including military, employment, or advanced training. Most importantly, the initial Governing Board policy statements were fashioned after a set of common goals for the decentralized system of American education crafted at the Charlottesville Summit, a meeting of the nation’s governors during the first Bush Administration (National Governors Association, 1991).
These initial policy definitions were sorely criticized by various groups; they were judged as extending far beyond the capabilities of what a cross-sectional survey such as NAEP can substantiate. Several evaluations examined the initial policy definitions and concluded that the predictive statements (e.g., at the Advanced level “...show readiness for rigorous college courses...”) could not be validated using NAEP data (Linn et al., 1991; Stufflebeam, Jaeger, and Scriven, 1991). Consequently, the PDs were revised in 1993. The newer versions were streamlined and had no predictive statements, were fully balanced in that they applied to all grades and subjects equally well, and tapped into the cognitive processes related to the levels.

The full achievement level policy statement has been modified several times over the years (including as late as August 2007) to align it with recent legislative changes. However, the revised PDs in figure 3 have retained their saliency over the past 15 years. They serve to ensure that the standards being set on all NAEP subjects reflect the Board’s expectations for students in grades 4, 8, and 12, thereby fulfilling the Board’s legislative charge as the policy body responsible for setting the NAEP achievement levels.

It is important to state that, although the Board sets the expectations and adopts the achievement levels as policy at the completion of the process, there is much that happens in between that the Board monitors, directs, and approves with the assistance of many content and technical experts, policymakers, and stakeholders. More will be said about each of these activities in later sections of this paper.

Performance Level Descriptions

Perhaps one of the most important contributions that the National Assessment has made to the standard-setting movement is the pivotal role of performance level descriptions in the standard-setting process. The use of ALDs in standard setting in general was not common in 1990. In fact, there were no ALDs in the 1990 NAEP initiative (Hambleton and Bourque, 1991). Panelists were required to translate PDs directly into cut scores on the NAEP scale without benefit of the intermediary steps of using grade- and subject matter-appropriate descriptions of content. However, starting in 1992, the use of ALDs became standard operating procedure.

The timing of ALD development has varied from cycle to cycle. In one case the ALDs were developed after the fact for reporting purposes (1990), in another case they were developed by panelists during the standard-setting process itself (1992), and in another case they were developed a second time after the Board made the decision to adopt cut scores significantly different from those recommended in the standard-setting process (1996). In subsequent cycles, preliminary ALDs (viewed as working descriptions) were developed during the assessment framework development process by a national consensus content panel. These working descriptions were expanded and refined by panelists during the standard-setting process, and the modified and finalized ALDs were then used to report NAEP performance on each assessment along with exemplar items from the assessment.


**Current Policy Definitions**

(Excerpt from the Minutes of the Achievement Levels Committee
November 24, 1993)

**Proficient.** This level represents solid academic performance for each grade assessed. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.

**Basic.** This level denotes partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at each grade.

**Advanced.** This level signifies superior performance beyond Proficient.
In the early cycles that used this approach (1994 and 1996), the standard-setting panels were responsible to refine the preliminary ALDs during the standard-setting process in order to align the final version of the ALDs with the exact content of the assessment, which involved deleting content from the ALDs that was not measured and adding content to the ALDs that was measured.

However, in 1998, the process for crafting final ALDs was removed from the standard-setting panels altogether (ACT, 1997). Before implementing the 1998 level-setting process, content panels were brought together to examine the public comments from a broad review of the preliminary ALDs. The review, conducted via the Internet and focus groups, allowed the content panels to finalize the ALDs prior to the standard setting. The Board-approved ALDs became inputs to the standard-setting process as givens, just as the frameworks and the assessment items are givens. This approach is still being used in 2009.

When developing the ALDs, content experts have access to a number of assessment documents, including the assessment framework, test and item specifications, and the generic PDs. It is important to note that the ALDs do not derive from the item pool directly but from these more global assessment documents. The ALDs may be checked later for alignment with the item pool. However, because the ALDs are derived from assessment frameworks, they have a durability that serves NAEP well. The frameworks are reviewed only about every 10 years, while the item pools are revised (at least partially) each time a subject is given on a 2- or 4-year cycle. The approach allows the achievement level definitions to be used for reporting NAEP over the entire life of the frameworks.

How did the use of ALDs come about? Taking a closer look at the 1990 and 1992 standard-setting efforts, the Board concluded that an incorrect chronology was being used. It seemed backwards to develop the cut scores first and then, as an afterthought, develop descriptions that aligned with the cut scores. As a result, in 1994 preliminary ALDs were crafted by the consensus panels that develop the assessment frameworks. Who better knew the content and what students should know and be able to do than the panels that recommended the content of what would be assessed? The preliminary ALDs were considered preliminary because they would serve to delimit the domain and the assessment content and could act as guides for the item writers. In addition, it was likely that some content identified in the preliminary ALDs might not be included (for one reason or another) in the final selection of test items on the assessment. Therefore, having preliminary ALDs provided some flexibility later on. Further, the preliminary ALDs served a very important role in the standard-setting process; that is, they communicated to the panelists the Board’s expectations for Basic, Proficient, and Advanced for each grade tested and in the particular content measured. Figure 4 displays the initial NAEP ALDs in grade 12 mathematics excerpted from the 2003 NAEP assessment framework.
### Basic
Twelfth-grade students performing at the Basic level should demonstrate procedural and conceptual knowledge in solving problems in the five NAEP content strands. Twelfth-grade students performing at the Basic level should be able to use estimation to verify solutions and determine the reasonableness of results as applied to real-world problems. Twelfth graders performing at the Basic level should recognize relationships presented in verbal, algebraic, tabular, and graphical forms, and demonstrate knowledge of geometric relationships and corresponding measurement skills.

They should be able to apply statistical reasoning in the organization and display of data and in reading tables and graphs. They should be able to generalize from patterns and examples in the areas of algebra, geometry, and statistics. At this level, they should use correct mathematical language and symbols to communicate mathematical relationships and reasoning processes and use calculators appropriately to solve problems.

### Proficient
Twelfth-grade students performing at the Proficient level should consistently integrate mathematical concepts and procedures into the solutions of more complex problems in the five NAEP content strands. Twelfth-grade students performing at the Proficient level should demonstrate an understanding of algebraic, statistical, geometric, and spatial reasoning. They should be able to perform algebraic operations involving polynomials, justify geometric relationships, and judge and defend the reasonableness of answers as applied to real-world situations. These students should be able to analyze and interpret data in tabular and graphical form; understand and use elements of the function concept in symbolic, graphical, and tabular form; and make conjectures, defend ideas, and give supporting examples.

### Advanced
Twelfth-grade students performing at the Advanced level should consistently demonstrate the integration of procedural and conceptual knowledge and the synthesis of ideas in the five NAEP content strands. Twelfth-grade students performing at the Advanced level should understand the function concept and be able to compare and apply the numeric, algebraic, and graphical properties of functions. They should apply their knowledge of algebra, geometry, and statistics to solve problems in more advanced areas of continuous and discrete mathematics.

They should be able to formulate generalizations and create models through probing examples and counter-examples. They should be able to communicate their mathematical reasoning through the clear, concise, and correct use of mathematical symbolism and logical thinking.
The use of ALDs (sometimes referred to as performance level descriptors (PLDs) in other settings) in standard setting has become de rigueur for most agencies today; it was almost unheard of before the National Assessment. Today, ALDs/PLDs are used with virtually all standard-setting methods (Hambleton and Pitoniak, 2006). Hambleton and others argue that PLDs are critical to interpreting the score results and providing evidence of procedural validity of the process (Hambleton, 2001; Mills and Jaeger, 1998). Furthermore, most states now use ALDs to develop standards on their own testing programs. The ALDs represent a range of performance across the score range. These descriptions help the panelists conceptualize and internalize the policy definitions before engaging in the tasks associated with the particular standard-setting method used. Without them, standard-setting panels are left to their own devices (and perhaps creativity) to recommend appropriate cut scores. Furthermore, because the panels are so diverse it is unlikely they would have the “common understanding” of the expectations of student performance that the policy board has in mind.

Finally, the ALDs are used to report assessment results. They provide a level of interpretation for the knowledge and skills that students within the levels know and can do. The ALDs, along with the exemplar items and the percentages of examinees at or above the cut scores, provide a reasonably broad picture of what the nation’s students in grades 4, 8, and 12 know and can do.

Another nuance of the ALDs worth mentioning here is the use of borderline descriptions in the level-setting process. As mentioned above, the ALDs are inputs to the process and, as such, they outline the content that students should know for those whose performance is in the designated score range. For example, the Proficient achievement level description outlines the content expectations for students whose performance is in the Proficient range; i.e., from the Proficient cut score up to the Advanced cut score. However, in training panelists to assess item content and recommend cut scores, it is necessary for them to think about the borderline performance of examinees; that is, what do students need to know to move from the Basic level up to the Proficient level? Thus, borderline descriptions are a subset of the ALDs and include the content necessary to move from a lower level up to a higher level.

In NAEP, the borderline descriptions are developed by the panels during the standard-setting process after extensive training, during which they have developed an understanding of the assessment, framework, item pools, policy definitions, and most especially the ALDs. The borderline descriptions are usually developed in grade-level groups and are working documents than can be refined during the process.¹¹

At the end of two decades of level setting, the Board seems to have mastered the process for developing the descriptions of achievement. The Board sets the policy definitions, which are given to consensus panels or independent content panels to operationalize in terms of specific grade levels and content areas. The draft statements are widely vetted with a variety of NAEP audiences, including other content specialists, stakeholders, NAEP users, and policymakers. The results are incorporated into a revised version of the ALDs that the Governing Board approves before the standard setting panels use them to develop their recommendations on cut scores and exemplar items. Finally, the whole package (i.e., recommended cut scores, descriptions, and exemplars) is brought to the policy Board for review before a final decision is made.
Methodology for Developing Standards

In the early 1990s, as the Governing Board was initiating the NAEP standard-setting process, the choice of methods was limited. Berk (1986) and others (Cizek and Bunch, 2007) have described some of these early procedures, including the Nedelsky (1954) method, the Ebel (1972) procedure, and the Angoff method (1971). While all three of these early methodologies were primarily designed for use with multiple-choice items, the Angoff method was the most researched in the literature of the 1970s and 1980s. So, at the time of NAGB’s initial decision to choose a methodology for NAEP standard setting, the recommendation from experts was to use Angoff. The method is a “judgment” method, meaning that panels examined items and item content and made a “judgment” about the probability of examinees answering the item correctly.\(^\text{12}\) A distinct advantage was that the Angoff method did not require any empirical data such as examinee performance on the items.\(^\text{13}\) Further, it was fairly straightforward to train panelists to complete the required tasks in the Angoff method; it was easily explained to standard-setting panelists (educators and noneducators alike) and could be adapted to accommodate multiple levels and multiple item formats.

By the 1992 standard setting in mathematics, the Board had contracted with ACT in Iowa City to implement the Board’s policy. ACT was responsible for developing ALDs; convening national samples of grade 4, 8, and 12 panels; implementing pilot and research studies; conducting the standard-setting meetings; providing recommendations to the Board; and producing all process and technical reports. Reckase (2000) and Loomis and Bourque (2001a) both provide comprehensive descriptions of the research conducted by ACT on the different methodologies explored during this period (1992 to 2000).

Between the 1992 and 1998 cycles, ACT developed standards on seven NAEP subjects, including mathematics, reading, science, writing, civics, U.S. history, and geography. All seven used a modified Angoff procedure to develop the achievement levels.\(^\text{14}\) The method used by NAEP was eventually modified to the extent that it took on a new descriptor and eventually was called the ACT/NAEP method. It was during this period that the ACT technical staff, with the advice of their Technical Advisory Committee on Standard Setting (TACSS), expanded and refined the standard-setting process in a number of ways.\(^\text{15,16}\)

First, training of panelists was standardized. That is, all grade-level panelists (4, 8, and 12) were trained in large group sessions by the coordinator of the session. All groups were exposed to the same training principles. Large group training sessions were complemented by grade-level group sessions moderated by a facilitator who was trained by the coordinator. The grade-level groups reinforced the large group training sessions and offered opportunities for panelists to ask questions and explore the tasks at hand. Standardization of training is desirable in standard setting just as it is in test administration. Results in either case should not be a function of who was involved. Minimizing the unintended effects of different facilitators is a laudable and necessary goal if replication across grades, content areas, etc. is to achieve common standards.

Second, panelists were trained and encouraged to internalize every aspect of the NAEP assessments (including the NAEP assessment framework, PDs, ALDs, and item formats being used in a particular assessment) before moving on to the standard-setting task. Customized
“practice exercises” were developed to assist in this process (e.g., panelists take a student-length version of the NAEP assessment or, in areas such as reading (1992), panelists review student-constructed response papers). It was critical for participants to understand the content aspects of the NAEP assessments and the expectations for the standards embodied in the PDs and ALDs. The better part of two days (of a five-day meeting) was spent at each standard-setting event to ensure that panelists had indeed internalized the salient content aspects of NAEP before proceeding.

Third, systematic feedback to panelists during the standard-setting process was designed to make the panelists better informed “judges” as they developed the recommended standards. The standard-setting process was viewed not as providing simply a professional opinion about the standards, but rather one’s professional judgment about the appropriate standards. It is believed that the better informed judge is the better standard setter. In 1992, the notion of feedback was somewhat novel in most standard setting, and was carefully crafted and cultivated as ACT improved on each NAEP cycle. Feedback was staged and provided judiciously so as to not influence unduly the judgments being made during the rating process. Such feedback examined inter- and intra-rater location data, rater consistency feedback, empirical data such as p-values, whole booklet feedback, and other useful “reality checks” for panelists to ground themselves. Panelists were always free to use (or not to use) the feedback data as they saw fit in making their judgments and recommendations.

Fourth, the process (not the panelists) was monitored from beginning to end using a series of self-report questionnaires completed by every panelist. The evaluations were used to improve future processes and to determine whether the panelists felt confident in the work they had completed. These process evaluations also served as evidence of procedural validity, asking a variety of questions about their understanding of the assessment content, the training provided, the methodology being used, the feedback provided, and their level of confidence in and satisfaction with their overall recommendations to the Board.

Standardized and comprehensive training, extensive use of feedback, and formal process evaluations were all modifications added to the original Angoff method. In addition, three cut scores were recommended and the probability of a correct response to an item was expanded from “0% or 100%,” as the original footnote suggested (Angoff, 1971), to 0 percent up to 100 percent and any probability in between.

Reflecting on the Legislation

The first four implementation issues described above—number of levels, policy definitions, performance level descriptors, and methodology—are not directly addressed by the legislation. The Board was initially, and still is, free to view these issues as operational and/or program policy issues, and has great latitude to formulate its decisions. However, that being said, the decisions should be made with advice from experts in the field, and full consideration of the professional guidelines for standard setting issued by the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education. The current set, issued as part of the Standards for Educational and Psychological Testing in 1999, provides guidance for (1) evidence of validity, e.g., the qualifications of experts
used in the test construction process (including formulating content, standards or scoring); (2) evidence of reliability, e.g., conditional standard errors of cut scores and repeated measures reliability; and (3) test interpretation, e.g., documentation of rationale and procedures for establishing cut scores. Each of these areas provides specific guidance for the Board and its contractors, and attempts have been made over the years to collect this kind of information during the standard-setting process and to align the standard-setting process with such guidelines.

Composition of the Standard-Setting Panels

The Governing Board’s first policy statement said that, “the panels be composed of individuals with expertise in the education of students of the ages and grades under consideration …[and] with knowledge of the typical subject area achievement…” (National Assessment Governing Board, 1990, p. 18). The approach was good; this was the acceptable standard in the field at the time. The problem was how do you find these panelists? The policy statement went on to say that, “Major national organizations will be contacted to recommend from among their members individuals who might serve on the panels…” (National Assessment Governing Board, 1990, p. 18). Two additional criteria included (1) continuity with the framework panels and (2) states participating in the new 1990 state-by-state assessments must be represented, especially at the eighth grade level.

Unfortunately, the 1990 effort resulted in panels that were too one-sided. Most participants were educators: classroom teachers, school administrators, and representatives of national education organizations. There were no business participants and no experts in the noneducation fields of mathematics. Because the education establishment was primarily represented, a tone was set at the meeting that was not conducive to setting standards. In reviewing that meeting, the Board came to the conclusion that the panels needed to be more diversified. To paraphrase one Board member at the time, setting standards on NAEP is just too important a task to involve only educators; we must involve others who also have a stake in the future of our country: employers, business and industry, and other stakeholders. With that in mind, the Board’s 1992 policy on panel composition required a distribution of 55 percent teacher-educators and 45 percent nonteachers, split between 15 percent nonteacher educators and 30 percent general public who are noneducators (American College Testing, 1992). In 1995, the Board policy was slightly modified to state that about two-thirds of the panels should be teachers and other educators, and one-third should be from the public/noneducator sector. These members are drawn from a national sampling frame and are broadly representative of the NAEP regions, types of communities, various ethnicities, and genders. In 1994, the notion of having broadly representative standard-setting panels was codified in the Goals 2000 Educate America law. The 1995 policy on panel composition is still being used in 2009 and is consistent with the current legislation.

Reporting Student Achievement Levels

One of the initial reasons for moving in the direction of a standards-based approach to NAEP was to provide better and more understandable information to users, including Congress, governors, state test users, policymakers at all levels, and the American public. Therefore, how
the NAEP data were reported became a paramount concern to the Governing Board. Two issues had to be dealt with: (1) how to transition from anchor levels to achievement levels and (2) how to report what students do know when the standards describe what students should know.

NAEP assessment results in the late 1980s were reported using anchor levels. Scale anchoring was developed by ETS in the early 1980s in an effort to improve the understandability of NAEP data. Previously, NAEP results were reported using fairly simplistic approaches such as item cluster reporting and sometimes domain reporting. But in general, these approaches were unsatisfactory. ETS’ New Design for a New Era (Messick, Beaton, and Lord, 1983) spelled out a new way of providing insight into the NAEP results called anchor levels. As conceptualized by ETS, the anchoring process involved:

- selecting a range of points on the score scale;
- constructing “item maps based on examinee performance;”
- identifying items that “anchor” around the selected points;
- describing the collection of anchored items at each of the selected points.

The five points selected for NAEP were the mean and one and two standard deviations above and below the mean. The cross-grade scale for each subject ranged from 0 to 500 (anchored at grade 8) with a mean of 250 and a standard deviation of 50, yielding the five points: 150, 200, 250, 300, and 350. Items were identified for each anchor point, and the descriptions of the points were crafted by content experts. For each subject there was one set of anchor levels that spanned the three grades at which NAEP tests. Anchor items were identified based on several criteria, including the percentage of students answering the items correctly at the selected points (80 percent), the percentage answering correctly at the next lower point (30 percent fewer), the percentage answering incorrectly at the next lower point (at least 50 percent), and the sample size on which the percentages are based (at least 100). The first criterion is one that can have a major impact on which items are identified as anchor items. Research has shown that using higher percentages (e.g., 80 percent) can yield anchor items and descriptions that are considerably easier in content, while using lower percentages (e.g., 65 or 50 percent) results in more difficult descriptions (Kolstadt et al., 1998). Easier items on the item maps result in easier descriptions; harder items on the item maps result in harder descriptions.

Clearly this new statistical approach was an improvement over the earlier methods, but it still remained difficult for policymakers to make decisions based on such data. Even if one knew that the mean mathematics performance of grade 8 students was 242, no one knew whether 242 was “good enough.” What was known was that the mean performance was slightly below the overall mean on the assessment, and the scale value could be described using the content of the anchor items (what students know and can do). However, even though the anchor levels described what students did know and can do, do they reflect what students should know and be able to do? In other words, was the performance of 242 good enough?

The Board was authorized to set performance levels and in 1990 it did so on the new mathematics framework. The Board preserved the older frameworks and trend lines (called long-term trends) and launched the new assessments with a new way of reporting their results—
achievement levels. The Board took on the legislative charge of answering the question, “How good is good enough?”

Transitioning from anchor levels to achievement levels has not been an easy journey. Initially, both anchor levels and achievement levels appeared together in the NAEP reports (NCES, 1994; American College Testing, 1995). The nuanced distinctions between the two were difficult for NAEP users and policymakers to grasp. One was statistically based; the other standards based—easier said than understood. One described what students know and can do (the current condition of education); the other was based on what students should know and be able to do (the goals of education). The anchor levels were statements of the status quo; the achievements levels were expectations, desired outcomes for performance goals on NAEP. This issue of can and should became a hotly debated topic in the mid-1990s. The author’s sense is that it has never been resolved completely. However, the most cogent argument toward resolution was framed by Reckase, who argued that “can” and “should” represent a tautology. When speaking of the achievement levels as goals to be reached, the terminology “should” is used; when reporting on the percentage of examinees whose performance on NAEP is within the achievement level categories, the terminology “can” is used.

This is where the legislative terminology came into play. The 1994 legislation used the term “developmental” to describe the levels during the transition from anchor levels to achievement levels. The 2001 legislation changed that language (perhaps moving it a bit forward by using the terminology “trial”). However, the decision regarding when the levels are no longer a “trial” still resides with the Commissioner, not the Board. The NAEP state trial assessments were dubbed “trial” for three cycles (1990, 1992, and 1994). By 1994, they were simply the NAEP state assessments. The same caution will probably hold for the Trial Urban District Assessments and that label could be removed in the foreseeable future. However, an evaluation of the levels by the NAE (1997) concluded that, “…the current achievement levels be abandoned by the end of the century…” A study by the National Academy of Sciences (NAS) (Pellegrino, Jones, and Mitchell, 1998) concurred. The NAS study, however, presented no new research on the issue, but depended on older studies to reach its conclusion. It seems unlikely that the label “trial” as it applies to achievement levels will be removed any time soon.

**Contextual Issues at the Start of the 21st Century**

As the Board moved into the first decade of the 21st century, the NAEP frameworks in most subjects were approaching 10 years in use. This was especially true for the oldest in the set, mathematics and reading. It had always been the Board’s intention and policy to renew and revitalize the frameworks about every 10 years, and as movement took place in the curriculum field. That is not to say that new fads became the impetus for change. However, in the decade of the 1990s there were new developments in learning theory, and some of those developments impacted how and what was being taught in the schools. National professional organizations were rethinking their earlier curricula scope and sequence (National Council for Teachers of Mathematics, 2000); states were completing new, or revising older, state content standards in an effort to meet NCLB requirements; and new demands were being articulated for postsecondary education and training. Therefore, taking a serious look at the NAEP frameworks by curriculum experts and others who comprise the NAEP national consensus approach seemed to be in order.
This brought to the foreground new issues to be solved for this new era of the achievement levels.

**Implementation Decisions: The New Century of Achievement Levels**

Early in the current decade, several implementation issues needed to be resolved if the achievement levels were going to maintain their currency, and especially if they were to be helpful to states as they meet the requirements of NCLB and look ahead to its anticipated reauthorization. Should there be new frameworks in the major subjects—reading and math—covered by NCLB? If the answer to that question is “yes,” then that would call for, at the very least, a review of the original achievement levels to see if they still would apply to a new framework. Second, if the original achievement levels cannot apply, then new ones could be adopted by the Board. However, that could disrupt trend lines. Can the original and new achievement levels be linked? Should there be changes in the Governing Board achievement levels policy based on what has been learned in the 1990s? For example, should the Governing Board continue to use the original ACT/NAEP methodology for developing the levels or should a different method be employed? We will explore each of these issues in turn.

**New Levels for New Frameworks?**

The 2000 and 2003 NAEP mathematics assessments continued the trend started in 1990 with the original mathematics framework developed under the Board’s guidance. Figure 4 (earlier in this paper) displays the ALDs for the original achievement levels in grade 12. However, in 2005 the Board developed and adopted a new framework for the grade 12 mathematics assessment. According to Board documents, this change was warranted because the grade 12 mathematics curriculum had become more challenging over the last decade. This movement was based partly on the international studies that reported less than stellar performance for U.S. students in both mathematics and science. In addition, curricula changes at the state and local levels had the effect of sequencing more difficult content at lower grade levels, consequently making grade 12 content more demanding and thus out of step with the grade 12 original NAEP framework.

Whether or not the old and new assessments were aligned enough to be placed on the same reporting scale, and whether the new assessment could be reported using the original achievement levels is an empirical question. Even though the statistical methodology of NAEP (item response theory) can accommodate minor shifts in item content, difficulty, and format, NAEP always undertakes an extensive empirical effort to undergird major decisions when important changes are made. In this case, the evidence showed that the two math assessments could not be closely aligned because of new content, changes in administration and block design, and different rules on calculator usage (National Assessment Governing Board, 2007). As a result, the Board developed new achievement levels for the 2005 grade 12 assessment in math. Figure 5 displays the ALDs for the new grade 12 assessment. The reader can compare the differences in the grade 12 ALDs and judge the degree of difference in content.
Figure 5
Grade 12 Mathematics ALDs
(Excerpt from 2005 Nation’s Report Card: 12th Grade Mathematics)

**Basic**
Twelfth-grade students performing at the *Basic* level should be able solve mathematical problems that require the direct application of concepts and procedures in familiar situations. For example, they should be able to perform computations with real numbers and estimate results of numerical calculations. These students should also be able to estimate, calculate, and compare measures and identify and compare properties of two- and three-dimensional figures, and solve simple problems using two-dimensional coordinate geometry. At this level, students should be able to identify the source of bias in a sample and make inferences from sample results; calculate, interpret, and use measures of central tendency; and compute simple probabilities. They should understand the use of variables, expressions, and equations to represent unknown quantities. They should be able to solve problems involving linear relations using tables, graphs, or symbols; and solve linear equations involving one variable.

**Proficient**
Students in the twelfth grade performing at the *Proficient* level should be able to select strategies to solve problems and integrate concepts and procedures. These students should be able to interpret an argument, justify a mathematical process, and make comparisons dealing with a wide variety of mathematical tasks. They should also be able to perform calculations involving similar figures including right triangle trigonometry. They should understand and apply properties of geometric figures and relationships between figures in two and three dimensions. Students at this level should select and use appropriate units of measure as they apply formulas to solve problems. Students performing at this level should be able to use measures of central tendency and variability of distributions to make decisions and predictions, calculate combinations and permutations to solve problems, and understand the use of normal distribution to describe real-world situations. Students performing at the *Proficient* level should be able to identify, manipulate, graph, and apply linear, quadratic, exponential, and inverse proportionality ($y = k/x$) functions; solve routine and one-routine problems involving functions expressed in algebraic, verbal, tabular, and graphical forms; and solve quadratic and rational equations in one variable and solve systems of linear equations.

**Advanced**
Twelfth-grade students performing at the *Advanced* level should demonstrate in-depth knowledge of mathematical concepts and procedures represented in the framework. They can integrate knowledge to solve complex problems and justify and explain their thinking. These students should be able to analyze, make and justify mathematical arguments, and communicate their ideas clearly. *Advanced* level students should be able to describe the intersections of geometric figures in two and three dimensions, and use vectors to represent velocity and direction. They should also be able to describe the impact of linear transformations and outliers of measures on central tendency and variability, anlyze predictions based on multiple data sets, and apply probability and statistical reasoning in more complex problems. Students performing at the *Advanced* level should be able to solve or interpret systems of inequalities, formulate a model for a complex situations (e.g., exponential growth and decay), and make inferences or predictions using the mathematical model.
Developing new achievement levels in grade 12 mathematics was not an easy decision, but it exemplifies the complexities ahead of the Board as they review, renew, and/or revise the current frameworks in nearly a dozen different NAEP subject areas. In 1989, the Board maintained the initial trend lines (from approximately 1971 to 1989) by preserving the Long-Term Trend assessment, still administered and reported separately from the main NAEP.21 If that approach is used again, there could be two long-term trend lines, 1971 to 1989 and 1990 to 2005. The difference this time is that so much is riding on the trend lines from 1990, because, the NAEP state assessments began that year and the NCLB accountability requirements started in 2001. Admittedly, NCLB accountability applies to three subjects (reading/language arts, mathematics, and science). However, currently, the 2005 mathematics results in grade 12 are reported on a single-grade scale, with no links to earlier mathematics assessments. This approach in other subjects and other grades may not be entirely suitable to moving the accountability embedded in NCLB forward. More recently, a second new mathematics framework was developed for use in the NAEP 2009 cycle along with a new framework in reading at all three grades.22 A proliferation of frameworks without a clear plan to develop a unified assessment program with meaningful achievement levels will likely not serve the National Assessment well.

To Link or Not to Link?

Equating the results of one assessment to another, or one grade to another, has been statistically possible for a number of years. However, the procedure is not without its difficulties and critics. Equating is as much an art as it is a science. Depending on the decisions made in the linking process, results can vary substantially. At the August 2007 Board meeting, the Board heard a presentation on NAEP trend line issues that included a discussion of linking, achievement levels, and scaling in the context of the 2009 reading assessment that will employ a new NAEP reading framework (National Assessment Governing Board, 2007).

Representatives from the Education Information Advisory Committee of the Council of Chief State School Officers, the NAEP Validity Studies Panel, and the NAEP Design and Analysis Committee (an advisory committee to the current NAEP contractor, ETS) presented findings from studies and discussions that each group had conducted on the issue.

The recommendation resulting from the panel presentation was that new achievement levels should be set based on the new 2009 reading framework. The question regarding whether or not to link the old and the new levels is still open and under study by NCES and others. Whatever is decided for the 2009 reading assessment must be carefully and cautiously considered, since it paves the way for dealing with other subject areas and sets a precedent for the future of the NAEP achievement levels.

Original Methodology or New Methodology?

By 2003, the new NCLB law was beginning to take hold. Many states had received approval by the U.S. Department of Education to move forward with their own assessments and to develop accountability standards aligned to their assessments. As required under the law, states began to develop performance standards using NAEP’s standards as a model. The law requires two higher levels (e.g., Proficient and Advanced) and at least one lower level in order to
report on the performance of the full distribution. Those labels are not required and, indeed, many states use quite different labels to describe student performance, (e.g., Levels I, II, III, IV, or other descriptors such as Low, Intermediate, High). Also, states use variations on the number of levels (from three to six levels) (Perie, 2008). A review of state performance standards shows that 12 states use a 5-level system, 29 use a 4-level system, 10 use a 3-level system, and 1 uses a 6-level system. All states with three or four levels have positioned the required “Proficient” at the second highest level. Of the 13 states that use 5 or 6 levels, 9 have positioned the required “Proficient” at the third highest level, i.e., three levels down from the top level, thus having the likely effect of depressing the definition of Proficient.

Table 1 summarizes the number of performance levels by state, and where Proficient is positioned in the distribution of performance. Further, the definition of Proficient can vary from state to state, and is not required to reflect the NAEP definition. Both of these aspects, the positioning and definition of Proficient, affect the relationship of NCLB and the achievement levels. This is no small matter and its resolution would go a long way to resolving the disparity between NAEP results for the states and the states’ performance on their approved NCLB assessments.

| Table 1 |
| State Performance Standards under NCLB* |
| No. of Levels |
| Position of Proficient | 3 levels | 4 levels | 5 levels | 6 levels |
| N = 10 | N = 29 | N = 12 | N = 1 |
| 2nd highest level | CO GA | AL AK | CA MO | RI |
| IN IA | AZ AR | OR VT |
| MD NJ | DC HI | |
| PR TN | ID IL |
| TX VA | KY ME | |
| MA MI | MS MT |
| MS MT | NB NE |
| NH NM | NY NC |
| NY NC | ND OK |
| ND OK | PA SC |
| PA SC | SD UT |
| SD UT | WA WI |
| WA WI | WY |
| 3rd highest level | |
| CT DE |
| CT DE |
| FL KS |
| LA MN |
| OH WV |

*Adapted from Perie (2008).

Notes:
Example of 3 levels – Does not meet standard, meets standard, exceeds standard.
Example of 4 levels – Far below Proficient, below Proficient, Proficient, Advanced.
Example of 5 levels – Levels 1, 2, 3, 4, 5, where Level 3 = Proficient.
In 2003, the Board awarded a contract to ACT to develop the achievement levels on the grade 12 mathematics assessment. At the November 2004 meeting, the Executive Director’s report contained a summary of the Board’s ongoing work, including the new contract. ACT had already completed a number of pilot studies to look at the viability of moving to a new methodology for NAEP. A Committee on Standards, Design, and Methodology (COSDAM) meeting prior to the November meeting fully discussed the pros and cons of such a move, and recommended the change from Angoff to the MAP Mark method of setting achievement levels (Schultz and Mitzel, n.d.). This method is a variant on the Bookmark method, used by many states, in which items are arrayed according to difficulty and judges pick a “passing” score. (Nellhaus, 2000).23

According to the Board meeting transcript, the following factors influenced the Board’s decision:

1. “The new NAGB framework for 12th grade mathematics is sufficiently different from the previous framework to require a new trend line;
2. The MAP Mark and item rating methods are likely to produce valid outcomes;
3. The MAP Mark approach is based on the bookmark method that is widely used by states in setting achievement levels; and
4. MAP Mark is less complex and easier to explain and defend than the modified Angoff method” (National Assessment Governing Board, 2004).

A review of other publicly available Governing Board documents provided no other operational or policy reasons for the shift from Angoff to MAP Mark. According to a staff member, the Board reviewed evidence for both the ACT/Governing Board method and the MAP Mark method and judged them to be very similar. The Board’s committee responsible for the achievement levels (COSDAM) was interested in the fact that the newer method could be implemented in a shorter time period (four days rather than five).24 This new method was subsequently used for the grade 12 Economics assessment in 2006. Table 2 displays the NAEP achievement levels results in the first year the levels were set for all NAEP assessments since 1992.
<table>
<thead>
<tr>
<th>Content/Initial Year</th>
<th>Cut Score and Percent at or Above</th>
<th>Reported on NAEP 0 to 500 Cross-Grade Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Basic</td>
<td>Proficient</td>
</tr>
<tr>
<td>Math/1992</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 4</td>
<td>214/59%</td>
<td>249/18%</td>
</tr>
<tr>
<td>Grade 8</td>
<td>262/58</td>
<td>299/21</td>
</tr>
<tr>
<td>Grade 12</td>
<td>288/64</td>
<td>336/15</td>
</tr>
<tr>
<td>Reading/1992</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 4</td>
<td>208/62</td>
<td>238/29</td>
</tr>
<tr>
<td>Grade 8</td>
<td>243/69</td>
<td>281/29</td>
</tr>
<tr>
<td>Grade 12</td>
<td>265/80</td>
<td>302/40</td>
</tr>
<tr>
<td>Geography/1994</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 4</td>
<td>187/70</td>
<td>240/22</td>
</tr>
<tr>
<td>Grade 8</td>
<td>242/71</td>
<td>282/28</td>
</tr>
<tr>
<td>Grade 12</td>
<td>270/70</td>
<td>305/27</td>
</tr>
<tr>
<td>U.S. History/1994</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 4</td>
<td>195/64</td>
<td>243/17</td>
</tr>
<tr>
<td>Grade 8</td>
<td>252/61</td>
<td>294/14</td>
</tr>
<tr>
<td>Grade 12</td>
<td>294/43</td>
<td>325/11</td>
</tr>
<tr>
<td>Science/1996</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 4</td>
<td>138/67</td>
<td>170/29</td>
</tr>
<tr>
<td>Grade 8</td>
<td>143/61</td>
<td>170/29</td>
</tr>
<tr>
<td>Grade 12</td>
<td>145/57</td>
<td>178/21</td>
</tr>
<tr>
<td>Civics/1998</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 4</td>
<td>136/69</td>
<td>177/23</td>
</tr>
<tr>
<td>Grade 8</td>
<td>134/70</td>
<td>178/22</td>
</tr>
<tr>
<td>Grade 12</td>
<td>139/65</td>
<td>174/26</td>
</tr>
<tr>
<td>Writing/1998</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 4</td>
<td>115/84</td>
<td>176/23</td>
</tr>
<tr>
<td>Grade 8</td>
<td>114/84</td>
<td>173/27</td>
</tr>
<tr>
<td>Grade 12</td>
<td>122/78</td>
<td>178/22</td>
</tr>
<tr>
<td>Math/2005</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 12</td>
<td>141/61</td>
<td>176/23</td>
</tr>
<tr>
<td>Economics/2006</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 12</td>
<td>123/79</td>
<td>160/42</td>
</tr>
</tbody>
</table>

Sources: Loomis and Bourque (2001b); Grigg, Donahue, and Dion (2007); Mead and Sandene (2007).
Impact of Achievement Levels Over the Last 20 Years

Where are we after 20 years of standards on NAEP? One thing is certain, the sky did not fall (as predicted by some naysayers). The Nation’s Report Card, while not on the New York Times bestseller list, certainly makes news headlines when the results are released. Education shows evidence of being somewhat better off than it was in 1989. The dialog has picked up speed about quality education and what we as a nation should expect from the system. Not that achievement levels can take all the credit, but they may have helped. The author believes that the Governing Board’s achievement levels can stand tall and take credit for a number of advances in the standards movement. The achievement levels:

- improve the form and use of NAEP;
- serve policy decisionmaking efforts at the local, state, and federal levels;
- serve as a model for state assessments under NCLB;
- improve the standard-setting enterprise.

Clearly, the levels have improved the form and use of NAEP. Although there will always be some concerns about the “clarity” of the levels, for most users they are far better than what existed before, namely, average scale scores with descriptions at the mean, the 84th percentile, etc. The levels have tried to answer the question, “How good is good enough?” We may not agree with the answer to the question any more than we agree with how clean our air should be or how “green” our automobiles should be. Nevertheless, we need the answers to those kinds of questions in order to measure our progress toward national and world goals. The same is true in education: we need the answer to “How good is good enough?” to measure our progress as a nation toward an educated citizenry.

The levels also serve a very important policy function at all levels of education. The importance of NAEP results that policymakers may use cannot be overstated. The move in 1990 to add the Trial State Assessments to the mix, and the more recent move to add the Trial Urban District Assessments as well, is a testimony to NAEP’s value in general and the achievement levels in particular. The levels have influenced legislation at the federal and state levels, they have been used to provide snapshots of academic performance by news outlets such as Education Week, and they are used by private foundations and think tanks whose mission is to keep the public informed on the condition of U.S. education. See, for example, the recent American Institutes for Research report on linking international mathematics performance with NAEP mathematics performance in urban school districts (Phillips and Dossey, 2008).

The levels were also invoked as a model for states in the NCLB legislation. States are required to move toward a standards-based approach in reporting the state’s progress in reaching the accountability goals of NCLB. They must set levels of performance similar to the achievement levels in that they span the distribution of performance. All states now have those levels, along with assessments and aligned content standards.

Finally, the Governing Board achievement levels have improved the standard-setting enterprise enormously. In 1989 there was a paucity of methods that could be used to set standards; no one had heard of PDs or ALDs; use of feedback during the standard-setting process
was slim to none; using broadly based panels was never done; and, in many instances, the process to set standards was done “behind closed doors.” All that has changed; in no small measure, that change is due to the work of the Governing Board.

More methods are available now than ever before—more than one-third of the states use PDs and all use ALDs; virtually all methods use feedback in the process; and, in many instances, the composition of panels has been expanded and the process is more open. In addition, the level of discourse in the professional literature has increased considerably.

What Does the Future Hold?

The National Assessment Governing Board is to be congratulated on this 20th anniversary for its work on the NAEP performance standards. Over the last 20 years, the Board has developed, defended, and disseminated the achievement levels. But—five Administrations and 10 congressional sessions later—the work is not done. Many major issues are still not resolved. And even though resources may be limited, it is important to keep in mind that progress can be measured only if the yardstick used for measuring is valid and reliable, and reflects the current best practice in measurement technology. Here is the author’s list of what should be done:

1. Work to remove “trial” from the next NAEP reauthorization.
2. Resolve the issues of trend line and single-grade scale.
3. Explore the possibility of linking new and old assessments, especially in the short-term trend.
4. Resolve the discrepancies in performance between NAEP and state assessment results, or explain them much better.
5. Explore frameworks and achievement levels in the context of 21st century skills, and build new frameworks in accordance with a deliberate plan that preserves the integrity of the NAEP program for the foreseeable future.
6. Mount a robust research agenda on new standard-setting methodologies and publish the results in the professional literature. Items on the agenda should include:
   a. Evidence of bias or lack thereof in the overall standard-setting results.
   b. Impact of various criteria for selecting anchor items and ordering them on the scale.
   c. Impact of item formats and methodology and their interactions on the levels.
   d. Validity data, validity data, and more validity data

Endnotes

1. Much of the information in this section is taken from Jones and Olkin (2004).
2. This thumbnail sketch of the early NAEP decades does not allow extensive coverage of the pressing domestic issues during this period. Suffice it to say that events such as the landmark 1954 Supreme Court decision on desegregating public schools, the development of the 1965
Great Society programs of President Johnson, passage of the 1965 Title I legislation, and the 1966 Coleman report on educational inequality also influenced the crafting of a national measure of educational progress.

3. The 1990 NAEP cycle was the first developed under the Board’s policy. This assessment cycle reflected some significant changes, including developing assessment frameworks through a national consensus process, moving from age-based sampling and reporting to grade-based sampling and reporting, preserving the first 20 years of NAEP assessments as the Long-Term Trend, and reporting NAEP performance in terms of the achievement levels rather than using statistical indicators of the national distribution (e.g., means and standard deviations). See Phillips et al. (1993) for more specific information on this topic.


5. For a more detailed discussion of the distinctions between anchor level and achievement levels, see Bourque (2007).

6. The three achievement levels have been used in all NAEP reports since 1992. However, there is also a “Below Basic” level that the Board does not view as a NAEP standard; it is included only to complete the reporting of the full distribution of student performance.

7. One must distinguish here between comparability of standards and comparability of results. The former is a policy requirement for NAEP. However, comparability of results cannot be guaranteed since achievement is the result of many factors, including curriculum emphasis, time allocations in schools, grade structure of state curricula, the percent of special needs participation in NAEP, and other issues beyond the control of NAEP.

8. A Foreword to the Achievement Levels Policy and Implementation Guidelines was added and adopted by the Board in August 2007; it explains and updates policy changes between 1990 and 2007.

9. This paper will refer to these as achievement level descriptions (ALDs), but in the standard setting literature they are commonly referred to as performance level descriptions (PLDs).

10. The 1992 standard setting did employ operationalized versions of the policy definitions. However, they were developed during the standard-setting meeting by three distinct grade-level groups and, as such, varied in sharpness of the language, degree of specificity, and format. These working versions were subsequently validated by an independent group to sharpen the language and to provide durability to the descriptions.

11. It should be noted that the use of borderline descriptions discussed here applies only to the original method of setting standards adopted by the Board in 1990 (the Angoff method). The method adopted more recently for 2005 grade 12 mathematics and 2006 economics does not employ borderline descriptions.
12. This approach is in contrast to other “norm-referenced” approaches in which targets or quotas are set without much regard for the content of the standards. See Hambleton and Pitoniak (2006), Plake (2005), and Goodman and Hambleton (2005) for further information.

13. Although data are not a requirement, the Angoff method (as well as most other methods) is almost always used with some data—even field test data—to provide feedback to panelists (a reality check) during the process.

14. The Angoff method as originally suggested by Angoff has almost always been accommodated (a.k.a. modified Angoff) to particular circumstances.

15. ACT technical staff included Robert Brennan and Mark Reckase in addition to other ACT technical staff who served on an internal Technical Advisory Team.

16. Initial members on TACSS included William Brown, South Carolina Department of Public Instruction; Robert Forsyth, University of Iowa; Ronald Hambleton, University of Massachusetts; Eugene Johnson, ETS; Michael Kane, University of Wisconsin; Brenda Lloyd, University of Virginia; and William Mehrens, Michigan State University.

17. Consequences data were not allowed to impact the process in any meaningful way until the 1998 cycle. Initially, the Board reserved the use of those data for itself; however, by 1998, the Board was convinced that the use of consequences data could improve the panelists’ judgments, and allowed its use in the final round of the process.

18. The purpose of this technical paper was not focused on scale anchoring or standard setting per se, but its results lend support to the statements here and their applicability to both approaches for reporting.

19. Full disclosure would admit that there is another interpretation. NCES, which is responsible for the administration of NAEP (but not policy) argues that as a statistical agency it should report only NAEP data, without any accompanying judgments about “how good is good enough.” And so “developmental” also has connotations of “use with caution.”

20. NCES did not mount any specific bridge studies to gather empirical data on the impact of the changes.

21. The Long-Term Trend assessment uses smaller age-based sample sizes, includes only a national assessment but no state assessments, and preserves the same exclusion rules that had been initially used in NAEP before 1990.

22. S. Loomis, personal communication (2009).

23. A comprehensive review of the literature pertaining to the Bookmark method can be found in Karantonis and Sireci (2006).

References


Reckase, M.D. (2000). *The evolution of the NAEP achievement level setting process: A summary of the research and development efforts conducted by ACT.* ACT: Iowa City, IA.


Discussion of Proposed Revised Policy Statement on NAEP Achievement Level Setting

For the past 1.5 years, COSDAM members have been discussing the need to update the 1995 Governing Board policy statement on Developing Student Performance Levels for NAEP. The Board’s formal response to the November 2016 evaluation of the NAEP achievement levels (attached) noted that several of the report recommendations would be addressed through a revision of the Board policy statement. In particular, the Board’s response stated that the updated policy will specify a process and timeline for conducting regularly recurring reviews of the achievement level descriptions (ALDs) and will be explicit about the conditions that necessitate consideration of a new standard setting. In addition, one of the planned activities for the implementation of the Strategic Vision is to consider new approaches to creating and updating the achievement level descriptions in the revision of the Board policy on achievement levels.

Given that the policy is over 20 years old, there was also a need to revisit the operational guidance for setting achievement levels to ensure that it reflects currently accepted best practices.

Several activities have informed the draft revised policy statement, including ongoing COSDAM discussion, input from technical experts in standard setting, and several research efforts. The March and May 2018 COSDAM meetings were entirely devoted to this topic. Full Board discussions on this topic were held in August 2017 and August 2018, and additional discussion and action are planned for the upcoming November 2018 meeting.

The table below summarizes both prior and planned upcoming activities:

<table>
<thead>
<tr>
<th>Activity</th>
<th>Timeline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial COSDAM discussions of planned activities and timeline</td>
<td>March – May 2017</td>
</tr>
<tr>
<td>Initial input from standard setting experts</td>
<td>March – April 2017</td>
</tr>
<tr>
<td>Initial full Board discussion about potential elements of policy revision</td>
<td>August 2017</td>
</tr>
<tr>
<td>Technical advisory panel to seek expert advice and debate on best practices in achievement level setting</td>
<td>January 2018</td>
</tr>
<tr>
<td>Literature review of best practices for creating and updating ALDs</td>
<td>February 2018</td>
</tr>
<tr>
<td>Technical memo on developing a validity argument for the NAEP achievement levels</td>
<td>February 2018</td>
</tr>
<tr>
<td>COSDAM discussion of using research to inform goals for policy revision</td>
<td>March 2018</td>
</tr>
<tr>
<td>COSDAM review and discussion of draft revised policy</td>
<td>May 2018</td>
</tr>
<tr>
<td>Joint COSDAM/R&amp;D discussion on communicating achievement levels</td>
<td>May 2018</td>
</tr>
<tr>
<td>COSDAM call to finalize a consensus draft for full Board discussion</td>
<td>June 2018</td>
</tr>
<tr>
<td>Technical advisory panel to seek expert advice on policy and procedures for reviewing and revising achievement level descriptions</td>
<td>July 2018</td>
</tr>
<tr>
<td>Full Board review and discussion of draft revised policy</td>
<td>August 2018</td>
</tr>
<tr>
<td>Public comment on draft revised policy</td>
<td>August 30 - October 15, 2018</td>
</tr>
<tr>
<td>Calls to discuss additional proposed revisions to draft policy</td>
<td>October 2018</td>
</tr>
<tr>
<td><strong>Planned discussions and full Board action on revised policy</strong></td>
<td><strong>November 2018</strong></td>
</tr>
</tbody>
</table>
Compared to the current (1995) policy on Developing Student Performance Levels for NAEP, the attached proposed revised policy statement reflects:

- Reorganization of principles, streamlining of language, minimization of redundancies
- Minor (non-substantive) edits to the NAEP policy definitions for clarity
- A change in terminology from Proficient to NAEP Proficient to better differentiate the NAEP achievement levels from other common uses of Basic, Proficient, Advanced
- A new principle on periodic review of achievement level descriptions and cut scores, prompted by the Board’s response to the evaluation of NAEP achievement levels
- A new principle to clarify participation of multiple stakeholders at various points throughout process
- A new principle to summarize the role of the Board
- Reference to an interpretative guide that would accompany the release of NAEP results and explain how the achievement levels should (and should not) be used
- Reference to multiple types of achievement level descriptions (ALDs), including reporting ALDs that would be created using empirical data and written in terms of what students do know and can do rather than what students should know and be able to do
- Clarification on the standard setting participants, in particular the non-educator group
- Additional details about the achievement level setting process, including some practices that have become institutionalized over time (e.g., the use of “impact data”)
- Removal of details on implementation directed to staff and contractors, which will instead be included in a “procedures manual”

The proposed policy revision is intended to be just one step towards implementing the Board’s formal response to the evaluation of NAEP achievement levels. Several other activities are underway and/or planned, and the overall status of all activities (both relevant to and distinct from the policy revision) will be discussed at the upcoming November COSDAM meeting.

Public Comment

Following the full Board discussion in August 2018, the Governing Board sought public comment on the proposed revised policy via a dedicated page on its website from August 30 – October 15, 2018: [https://www.nagb.gov/news-and-events/calendar/public-comment-on-als-policy.html](https://www.nagb.gov/news-and-events/calendar/public-comment-on-als-policy.html). Notices were posted in the federal register¹ on September 10th and October 2nd.

Seventy-three comments were received and have been compiled (see attached). Some of the major themes noted in the comments are as follows:

¹ The initial public comment period was intended to take place from August 30 – September 30, 2018. However, because the federal register notice was not actually published until September 10, the deadline was extended until October 15th. The typical timeframe for public comments from federal register notices is 30-45 days.
• Discussion of the value or lack thereof of the NAEP achievement levels in general
• Discussion of the policy labels and/or definitions, including:
  o Concerns that *NAEP Basic, NAEP Proficient, and NAEP Advanced* will not result in achievement levels that are “reasonable, valid, and informative to the public”
  o Praise for the proposal to add NAEP in front of *Basic, Proficient, and Advanced* to better distinguish from other uses of these terms, and support for the reasonableness of the NAEP achievement levels in comparison with other rigorous assessments
  o Concern that the original labels of *Basic, Proficient, and Advanced* should be retained without adding NAEP in front
• Discussion of the rigor of the NAEP achievement levels, including both the importance of maintaining high standards and calls for lowering standards
• Praise by technical experts for reflecting many current best practices in standard setting
• Concerns and suggestions for better communicating the meaning of achievement levels and achievement level results to various stakeholders
• Suggestions to connect the NAEP achievement levels with external benchmarks of success, either for validation purposes or as an alternative method of setting standards
• Discussion of the Board’s commitment to periodically review the NAEP achievement level descriptions and cut scores, including endorsement of this idea and concerns that they should not be changed unless absolutely necessary
• Suggestions for additional procedural details to be included in the policy
• A suggestion to include information about additional work to examine alignment between NAEP frameworks and state content standards

COSDAM held two teleconferences in mid-October to discuss the comments received and potential additional revisions to the policy. The attached proposed revised policy document reflects the outcomes from those calls; both a clean copy and a “tracked changes” version (in comparison to the version of the policy discussed at the August 2018 Board meeting and submitted for public comment) are included.
National Assessment Governing Board’s Response to the National Academies of Sciences, Engineering, and Medicine 2016 Evaluation of NAEP Achievement Levels

Legislative Authority

Pursuant to the National Assessment of Educational Progress (NAEP) legislation (Public Law 107-279), the National Assessment Governing Board (hereafter the Governing Board) is pleased to have this opportunity to apprise the Secretary of Education and the Congress of the Governing Board response to the recommendations of the National Academies of Sciences, Engineering, and Medicine evaluation of the NAEP achievement levels for mathematics and reading (Edley & Koenig, 2016).

The cited legislation charges the Governing Board with the authority and responsibility to “develop appropriate student achievement levels for each grade or age in each subject area to be tested.” The legislation also states that “such levels shall be determined by... a national consensus approach; used on a trial basis until the Commissioner for Education Statistics determines, as a result of an evaluation under subsection (f), that such levels are reasonable, valid, and informative to the public; ... [and] shall be updated as appropriate by the National Assessment Governing Board in consultation with the Commissioner for Education Statistics” (Public Law 107-279).

Background

NAEP is the largest nationally representative and continuing assessment of what our nation’s elementary and secondary students know and can do. Since 1969, NAEP has been the country’s foremost resource for measuring student progress and identifying differences in student achievement across student subgroups. In a time of changing state standards and assessments, NAEP serves as a trusted resource for parents, teachers, principals, policymakers, and researchers to compare student achievement across states and select large urban districts. NAEP results allow the nation to understand where more work must be done to improve learning among all students.

For 25 years, the NAEP achievement levels (Basic, Proficient, and Advanced) have been a signature feature of NAEP results. While scale scores provide information about student achievement over time and across student groups, achievement levels reflect the extent to which student performance is “good enough,” in each subject and grade, relative to aspirational goals.
Since the Governing Board began setting standards in the early 1990s, achievement levels have become a standard part of score reporting for many other assessment programs in the US and abroad.

Governing Board Response

Overview

The Governing Board appreciates the thorough, deliberative process undertaken over the past two years by the National Academies of Science, Engineering, and Medicine and the expert members of the Committee on the Evaluation of NAEP Achievement Levels for Mathematics and Reading. The Governing Board is pleased that the report concludes that the achievement levels are a meaningful and important part of NAEP reporting. The report states that, “during their 24 years [the achievement levels] have acquired meaning for NAEP’s various audiences and stakeholders; they serve as stable benchmarks for monitoring achievement trends, and they are widely used to inform public discourse and policy decisions. Users regard them as a regular, permanent feature of the NAEP reports” (Edley & Koenig, 2016; page Sum-8). The Governing Board has reviewed the seven recommendations presented in the report and finds them reasonable and thoughtful. The report will inform the Board’s future efforts to set achievement levels and communicate the meaning of NAEP Basic, Proficient, and Advanced. The recommendations intersect with two Governing Board documents, the Strategic Vision and the achievement levels policy, described here.

On November 18, 2016, the Governing Board adopted a Strategic Vision (https://www.nagb.org/content/nagb/assets/documents/newsroom/press-releases/2016/nagb-strategic-vision.pdf) to guide the work of the Board through 2020, with an emphasis on innovating to enhance NAEP’s form and content and expanding NAEP’s dissemination and use. The Strategic Vision answers the question, “How can NAEP provide information about how our students are doing in the most innovative, informative, and impactful ways?” The Governing Board is pleased that several of the report recommendations are consistent with the Board’s own vision. The Governing Board is committed to measuring the progress of our nation’s students toward their acquisition of academic knowledge, skills, and abilities relevant to this contemporary era.

The Governing Board’s approach to setting achievement levels is articulated in a policy statement, “Developing Student Performance Levels for the National Assessment of Educational Progress” (https://www.nagb.org/content/nagb/assets/documents/policies/developing-student-performance.pdf). The policy was first adopted in 1990 and was subsequently revised in 1995,
with minor wording changes made in 2007. The report motivates the revision of this policy, to add clarity and intentionality to the setting and communication of NAEP achievement levels.

The seven recommendations and the Governing Board response comprise a significant research and outreach trajectory that the Governing Board can pursue over several years in conjunction with key partners. The Governing Board will implement these responses within resource constraints and in conjunction with the priorities of the Strategic Vision.

**Evaluating the Alignment of NAEP Achievement Level Descriptors**

**Recommendation #1:** Alignment among the frameworks, the item pools, the achievement-level descriptors, and the cut scores is fundamental to the validity of inferences about student achievement. In 2009, alignment was evaluated for all grades in reading and for grade 12 in mathematics, and changes were made to the achievement-level descriptors, as needed. Similar research is needed to evaluate alignment for the grade 4 and grade 8 mathematics assessments and to revise them as needed to ensure that they represent the knowledge and skills of students at each achievement level. Moreover, additional work to verify alignment for grade 4 reading and grade 12 mathematics is needed.

The report’s primary recommendation is to evaluate the alignment, and revise if needed, the achievement level descriptors for NAEP mathematics and reading assessments in grades 4, 8, and 12. The Governing Board intends to issue a procurement for conducting studies to achieve this goal. The Governing Board has periodically conducted studies to evaluate whether the achievement level descriptors in a given subject should be revised, based on their alignment with the NAEP framework, item pool, and cut scores. The Governing Board agrees that this is a good time to ensure that current NAEP mathematics and reading achievement level descriptors align with the knowledge and skills of students in each achievement level category. In conjunction with the response to Recommendation #3, the updated Board policy on NAEP achievement levels will address the larger issue of specifying a process and timeline for conducting regular recurring reviews of the achievement level descriptions in all subjects and grades.

The Governing Board agrees strongly with the recommendation that, while evaluating alignment of achievement level descriptors is timely, it is not necessary to consider changing the cut scores or beginning a new trend line at this time. The NAEP assessments are transitioning from paper-based to digital assessments in 2017, and current efforts are focused on ensuring comparability between 2015 and 2017 scores. The Governing Board articulated this in the 2015 Resolution on Maintaining NAEP Trends with the Transition to Digital-Based Assessments (https://www.nagb.org/content/nagb/assets/documents/policies/resolution-on-trend-and-dba.pdf).

**Recommendation #2:** Once satisfactory alignment among the frameworks, the item pools, the achievement-level descriptors, and the cut scores in NAEP mathematics and reading has been
demonstrated, their designation as trial should be discontinued. This work should be completed and the results evaluated as stipulated by law: (20 U.S. Code 9622: National Assessment of Educational Progress: https://www.law.cornell.edu/uscode/text/20/9622 [September 2016]).

Ultimately, the Commissioner of Education Statistics is responsible for determining whether the “trial” designation is removed. The Governing Board is committed to providing the Commissioner with the information needed to make this determination in an expedient manner.

**Regular Recurring Reviews of the Achievement Level Descriptors**

**Recommendation #3:** To maintain the validity and usefulness of achievement levels, there should be regular recurring reviews of the achievement-level descriptors, with updates as needed, to ensure they reflect both the frameworks and the incorporation of those frameworks in NAEP assessments.

The Board’s current policy on NAEP achievement levels contains several principles and guidelines for setting achievement levels but does not address issues related to the continued use or reporting of achievement levels many years after they were established. The revised policy will seek to address this gap by including a statement of periodicity for conducting regular recurring reviews of the achievement level descriptors, with updates as needed, as called for in this recommendation. The Governing Board agrees that it is important to articulate a process and timeline for conducting regular reviews of the achievement level descriptors rather than performing such reviews on an ad hoc basis.

**Relationships Between NAEP Achievement Levels and External Measures**

**Recommendation #4:** Research is needed on the relationships between the NAEP achievement levels and concurrent or future performance on measures external to NAEP. Like the research that led to setting scale scores that represent academic preparedness for college, new research should focus on other measures of future performance, such as being on track for a college-ready high school diploma for 8th-grade students and readiness for middle school for 4th-grade students.

In addition to the extensive work that the Governing Board has conducted at grade 12 to relate NAEP mathematics and reading results to academic preparedness for college, the Governing Board has begun research at grade 8 with statistical linking studies of NAEP mathematics and reading and the ACT Explore assessments in those subjects. This work was published while the evaluation was in process and was not included in the Committee’s deliberations. Additional studies in NAEP mathematics and reading at grades 4 and 8 are beginning under contract to the National Center for Education Statistics (NCES). The Governing Board’s Strategic Vision includes an explicit goal to increase opportunities for connecting NAEP to other national and
international assessments and data. Just as the Board’s previous research related grade 12 NAEP results in mathematics and reading to students’ academic preparedness for college, the Governing Board anticipates that additional linkages with external measures will help connect the NAEP achievement levels and scale scores to other meaningful real-world indicators of current and future performance.

**Interpretations and Uses of NAEP Achievement Levels**

**Recommendation #5:** Research is needed to articulate the intended interpretations and uses of the achievement levels and collect validity evidence to support these interpretations and uses. In addition, research to identify the actual interpretations and uses commonly made by NAEP’s various audiences and evaluate the validity of each of them. This information should be communicated to users with clear guidance on substantiated and unsubstantiated interpretations.

The Governing Board’s Strategic Vision emphasizes improving the use and dissemination of NAEP results, and the Board’s work in this area will include achievement levels. The Governing Board recognizes that clarity and meaning of NAEP achievement levels (and scale scores) are of utmost importance. The Governing Board will issue a procurement to conduct research to better understand how various audiences have used and interpreted NAEP results (including achievement levels). The Governing Board will work collaboratively with NCES to provide further guidance and outreach about appropriate and inappropriate uses of NAEP achievement levels.

**Guidance for Inferences Made with Achievement Levels versus Scale Scores**

**Recommendation #6:** Guidance is needed to help users determine inferences that are best made with achievement levels and those best made with scale score statistics. Such guidance should be incorporated in every report that includes achievement levels.

The Governing Board understands that improper uses of achievement level statistics are widespread in the public domain and extend far beyond the use of NAEP data. Reports by the Governing Board and NCES have modeled appropriate use of NAEP data and will continue to do so. This recommendation is also consistent with the goal of the Strategic Vision to improve the dissemination and use of NAEP results. The Governing Board will continue to work with NCES and follow current research to provide guidance about inferences that are best made with achievement levels and those best made with scale score statistics.
Regular Cycle for Considering Desirability of Conducting a New Standard Setting

Recommendation #7: NAEP should implement a regular cycle for considering the desirability of conducting a new standard setting. Factors to consider include, but are not limited to: substantive changes in the constructs, item types, or frameworks; innovations in the modality for administering assessments; advances in standard setting methodologies; and changes in the policy environment for using NAEP results. These factors should be weighed against the downsides of interrupting the trend data and information.

When the Board’s achievement levels policy was first created and revised in the 1990s, the Board was setting standards in each subject and grade for the first time and had not yet considered the need or timeline for re-setting standards. To address this recommendation, the Governing Board will update the policy to be more explicit about conditions that require a new standard setting.

Board’s Commitment

The Governing Board remains committed to its congressional mandate to set “appropriate student achievement levels” for the National Assessment of Educational Progress. The Board appreciates the report’s affirmation that NAEP achievement levels have been set thoughtfully and carefully, consistent with professional guidelines for standard setting, and based on extensive technical advice from respected psychometricians and measurement specialists. The Board also takes seriously the charge to develop the current achievement levels through a national consensus approach, involving large numbers of knowledgeable teachers, curriculum specialists, business leaders, and members of the general public throughout the process. This is only fitting given the Governing Board’s own congressionally mandated membership that explicitly includes representatives from these stakeholder groups.

The Governing Board remains committed to improving the process of setting and communicating achievement levels. The Governing Board is grateful for the report recommendations that will advance these aims.

Reference

National Assessment Governing Board

Developing Student Achievement Levels for the National Assessment of Educational Progress

Policy Statement

It is the policy of the National Assessment Governing Board to conduct a comprehensive, inclusive, and deliberative process to develop and review student achievement levels for the National Assessment of Educational Progress (NAEP).\(^1\) Achievement levels consist of general policy definitions for the *NAEP Basic*, *NAEP Proficient*, and *NAEP Advanced* levels, specific achievement level descriptions (ALDs) for each assessment, cut scores that demarcate adjacent levels, and exemplar items or tasks that illustrate performance at each level. This process shall be conducted according to widely accepted professional standards, to produce results that are reasonable, useful, and informative to the public.

The Governing Board, through its Committee on Standards, Design and Methodology (COSDAM), shall monitor the development and review of student achievement levels to ensure that the final Governing Board-adopted achievement level descriptions, cut scores, and exemplars comply with all principles of this policy.

The achievement level setting process shall be carried out by contractors selected through a competitive bidding process. The process shall be managed in a technically sound, efficient, cost-effective manner, and shall be completed in a timely fashion.

Introduction

Since its creation by Congress in 1988, the Governing Board has been responsible for developing student achievement levels for NAEP assessments. The Governing Board has carried out this important statutory responsibility by engaging with a broad spectrum of stakeholders to develop student achievement levels.

---

\(^1\) According to current NAEP legislation, the Governing Board shall develop achievement levels for all NAEP assessments except for the Long-Term Trend assessment.
Under provisions of the National Assessment of Educational Progress Authorization Act of 2002 (P.L. 107-279), Congress authorized the Governing Board to, develop, “achievement levels that are consistent with relevant widely accepted professional assessment standards and based on the appropriate level of subject matter knowledge” (Section 303(e)(2)(A)(i)(II).

Given this mandate, the Governing Board must ensure that all achievement level setting processes align with current best practices in standard setting, and that appropriate validity evidence is collected and documented to support the intended uses and interpretations of NAEP achievement levels.

The Governing Board has established the following policy definitions for the NAEP achievement levels, as expectations of what students should know and be able to do. They shall be consistent across all assessments in which achievement levels are set.

**NAEP Basic**

This level denotes partial mastery of prerequisite knowledge and skills that are fundamental for performance at the NAEP Proficient level.

**NAEP Proficient**

This level represents solid academic performance for each NAEP assessment. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real world situations, and analytical skills appropriate to the subject matter.

**NAEP Advanced**

This level signifies superior performance beyond NAEP Proficient.

The Governing Board engages multiple stakeholders throughout the achievement level setting process, including:

- Teachers
- Curriculum Experts
- Content Experts
- Assessment Specialists
- State Administrators
- Local School Administrators
- Policymakers
- Business Representatives
- Parents
- Users of Assessment Data
- Researchers and Technical Experts
- Members of the Public

This policy also complies with the documents listed below which express widely accepted technical and professional standards for achievement level setting. These standards reflect the agreement of recognized experts in the field, as well as the policy positions of major professional and technical associations concerned with educational testing.
In conjunction with this policy the Board shall maintain a procedures manual to establish and document additional details about how this policy is to be implemented. As professional standards evolve and new consensus documents are released, this policy and the procedures manual shall be updated to the extent that new professional standards require. Resources for this purpose shall include, but not be limited to the following:


Principles for Setting Achievement Levels

Principle 1: Elements of Achievement Levels

Principle 2: Development of Achievement Level Recommendations

Principle 3: Validation and Reporting of Achievement Level Results

Principle 4: Periodic Review of Achievement Levels

Principle 5: Stakeholder Input

Principle 6: Role of the Governing Board
Principle 1: Elements of Achievement Levels

The Governing Board is responsible for developing student achievement levels for each NAEP assessment. Achievement levels for each NAEP assessment consist of content achievement level descriptions (ALDs), cut scores that demarcate adjacent levels, and exemplar items or tasks that illustrate performance at each level.

a) Content achievement level descriptions (ALDs) translate the policy definitions into specific expectations about student knowledge and skills in a particular content area, at each achievement level, for each subject and grade. Content ALDs provide descriptions of specific expected knowledge, skills, or abilities of students performing at each achievement level. Content ALDs reflect the range of performance that items and tasks should measure. During the achievement level setting process, the purpose of content ALDs is to provide consistency and specificity for panelist interpretations of policy definitions for a given assessment. During reporting, content ALDs communicate the specific knowledge and skills represented by *NAEP Basic*, *NAEP Proficient*, and *NAEP Advanced* for a given assessment.

b) Cut scores mark the minimum threshold score, the lower bound, for each achievement level. Performance within a given achievement level begins at the cut score for that level and ends just below the cut score for the successive achievement level.

c) Exemplar items or tasks, including student responses, illustrate student performance within each of the achievement levels. They provide specific examples to help the public better understand what students in each achievement level know and can do.

Principle 2: Development of Achievement Level Recommendations

The Governing Board shall develop student achievement levels for NAEP, consistent with relevant widely accepted professional assessment standards, based on the appropriate level of subject matter knowledge.

a) A Design Document shall be developed at the beginning of the achievement level setting process, to describe in detail the scope of the achievement level setting project being undertaken, including but not limited to all planned materials, procedures, and analyses needed for the project. The Design Document shall be posted for public review with sufficient time to allow for a response from those who wish to provide one.

b) The development of content achievement level descriptions (ALDs) shall be completed initially through the process that develops the assessment frameworks. (See the Governing Board Policy on Framework Development for additional details). The Board may then review and revise content ALDs to advance the purposes they serve, whether that is guiding an achievement level setting or informing the public about the meaning of achievement levels. Whether revised or not, the ALDs that guide
achievement level setting shall be articulated in terms of what students should know and be able to do. There shall be no content ALDs developed for performance below the NAEP Basic level.

c) An achievement-level setting panel of subject matter experts shall be convened to recommend achievement level cut scores and exemplars.

i. Each panel shall reflect diversity in terms of gender, race/ethnicity, region of the country, urbanicity, and experience with students with disabilities and English language learners. To ensure that they are qualified to make the judgments required by the achievement level setting process, individual panel members shall have expertise and experience in the specific content area in which the levels are being developed, expertise and experience in the education of students at the grade under consideration, and a general knowledge of assessment, curriculum, and student performance.

ii. Each panel shall include teachers, non-teacher educators, and other interested members of the general public with relevant educational background and experience. Teachers shall comprise the majority of the panel, with non-teacher educators (e.g., curriculum directors, academic coaches, principals) accounting for no more than half the number of teachers. The remaining panelists shall be non-educators who represent the perspectives of additional stakeholders representing the general public, including parents, researchers, and employers.

iii. The size of the panels shall reflect best practice in standard setting and be operationally feasible while being large enough to allow for split panels. Most NAEP achievement level settings have historically included approximately 20–30 panelists per grade, divided into two comparable groups.

d) Panelists shall receive training on all aspects of the achievement levels setting process to ensure that panelists are well-prepared to perform the achievement level setting tasks required of them. Panelists shall be instructed that their role is to make achievement level recommendations to the Governing Board. Training shall include but not be limited to: the purpose and significance of setting achievement levels for NAEP; the NAEP assessment framework for the given subject area; and administration of a sample assessment under NAEP-like conditions that students experience. It is important for panelists to arrive at a common conceptualization of NAEP Basic, NAEP Proficient, and NAEP Advanced based on the content ALDs. Panelists shall be trained on each element of the judgmental task they perform, including the selection of exemplar items. They should be led by capable content facilitators (who are content experts and have previous experience with achievement level setting) and process facilitators (who have background in standard setting and experience leading panelists through the achievement level setting process). Facilitators shall take a neutral stance and not attempt to influence panelist judgments.
e) The achievement level setting method that generates cut score recommendations shall have a solid research base and be appropriate for the content area, item types, number of items, scoring rubrics, and mode of administration, as applicable.

f) Evaluations shall be administered to panelists throughout the achievement level setting process, in accordance with current best practices. Evaluations shall be part of every major component of the process, and panelists shall be asked to confirm their readiness for performing their tasks. Evaluation data may be used for formative purposes (to improve training and procedures in future meetings); summative purposes (to evaluate how well the process was conducted and provide procedural validity evidence); and to inform the Governing Board of any relevant information that could be useful when considering cut score recommendations. The panelists shall have an opportunity to indicate to the Board whether they believe the recommended cut scores are reasonable.

g) In accordance with current best practices, feedback shall be provided to panelists, including “impact data” (i.e., the implications of their selected cut scores on the reported percentages of students at or above each achievement level).

h) The process shall consist of at least two achievement level setting meetings with distinct groups of panelists, a pilot study, and an operational meeting. The purpose of the pilot study is to conduct a full “dress rehearsal” of the operational meeting, including but not limited to: an opportunity to try out materials, training procedures, collection of panelist judgments, feedback given to panelists through the process, software used to conduct analyses, meeting logistics, and other essential elements of the process. The pilot study may result in minor changes to the procedures, as well as major changes that would need additional study before being implemented in an operational meeting. The pilot study provides an opportunity for procedural validity evidence and to improve the operational meeting. At the discretion of the Governing Board, other smaller-scale studies may be conducted prior to the pilot study or in response to issues raised by the pilot study. The criteria in Principle 2a apply to panelists of both meetings.

i) The Governing Board shall ensure that a Technical Advisory Committee on Standard Setting (TACSS) is convened to provide technical advice on all achievement level setting activities. Technical advice provided by standard setting experts throughout the project is intended to ensure that all procedures, materials, and reports are carried out in accordance with current best practices, providing additional validity evidence for the process and results. The Board or its contractor may also seek technical advice from other groups as appropriate, including NCES and the larger measurement community (e.g., the National Council on Measurement in Education).

j) All aspects of the procedures shall have documentation as evidence of the appropriateness of the procedures and results. This evidence shall be made available to the Board by the time of deliberations about the achievement levels. A summary of the evidence shall be available to the public when the achievement level results are reported.
k) Sample items and student responses known as exemplars shall be chosen from the pool of released items for the current NAEP assessment to reflect performance in the NAEP Basic, NAEP Proficient, and NAEP Advanced regions of the scale. The use of exemplars is intended to help the public better understand what performance in each achievement level represents for each subject and grade. When possible, exemplars may also be chosen that reflect performance at threshold scores. The collection of exemplars shall reflect the content found in the achievement level descriptions and the range of item formats on the assessment.

l) The outcomes from the achievement level setting panel meetings (recommended cut scores, exemplars, and ALDs for use in reporting) shall be forwarded to the Board for their consideration.

Principle 3: Validation and Reporting of Achievement Level Results

The achievement level setting process shall produce results that have validity evidence for the intended uses and interpretations and are informative to policy makers, educators, and the public.

a) Professional testing standards require evidence to support the intended interpretations and uses of test scores. Among the sources of evidence supporting the validity of test scores is evidence bearing on the standard setting process and results. Standard setting is necessarily judgmental, and the Board shall examine and consider available evidence about the procedural integrity of the achievement level setting process, the reasonableness of results, and other evidence in order to support intended uses and interpretations.

b) The Board shall examine and consider all evidence related to validity of the achievement level setting activities. These data shall include, but not be limited to: procedural evidence such as training, materials and panelist evaluation data; reliability evidence such as consistency across panelist type, subpanels, rounds, and meetings, if appropriate; and external comparisons to other similar assessments, if appropriate, with necessary caveats. The results from validation efforts shall be made available to the Board in a timely manner so that the Board has access to as much validation data as possible as it considers the recommendations regarding the final levels.

c) NAEP achievement levels are intended to estimate the percentage of students (overall and for selected student groups) in each achievement level category, for the nation, and for states and trial urban districts (TUDAs) for some assessments. NAEP is prohibited by law from reporting any results for individual students or schools.

d) In describing student performance using the achievement levels, terms such as “students performing at the NAEP Basic level” or “students performing at the NAEP Proficient level” are preferred over “Basic students” or “Proficient students”. The former implies that students have mastery of particular content represented by the achievement levels, while the latter implies an inherent
characteristic of individual students.

e) In reporting the results of NAEP, the three achievement levels of *NAEP Basic*, *NAEP Proficient*, and *NAEP Advanced* refer to the three regions of the NAEP scale at and above each respective cut score. The remaining region that falls below the *NAEP Basic* cut score shall be identified as “below *NAEP Basic*” when a descriptor is necessary.

f) In describing the *NAEP Proficient* level, reports shall emphasize that the policy definition is not intended to reflect “grade level” performance expectations, which are typically defined normatively and can vary widely by state and over time. *NAEP Proficient* may convey a different meaning from other uses of the term “proficient” in common terminology or in reference to other assessments.

g) To facilitate valid uses of ALDs for the purpose of reporting, the Board shall ensure that the descriptions of performance for the achievement levels reflect what the empirical data reveal about the knowledge and skills demonstrated by students in that score range. To develop ALDs for reporting, following the achievement level setting the Board shall revisit and may revise content ALDs to ensure that they are consistent with empirical evidence of student performance. In particular, these “Reporting ALDs” chosen to illustrate the knowledge and skills demonstrated at different achievement levels shall be written to incorporate empirical data from student performance. Reporting ALDs shall describe what students at each level *do* know and *can* do rather than what they *should* know and *should* be able to do.

h) An interpretative guide shall accompany NAEP reports, including specific examples of appropriate and inappropriate interpretations and uses of the results.

**Principle 4: Periodic Review of Achievement Levels**

*Periodic reviews of existing achievement levels shall determine whether new achievement level descriptions and/or cut scores are needed to continue valid and reliable measurement of current student performance and trends over time.*

a) At least once every 10 years or 3 administrations of an assessment, whichever comes later, the Governing Board, through its Committee on Standards, Design and Methodology (COSDAM), shall review the alignment between the content ALDs and items, based on empirical data from recent administrations of NAEP assessments. In its review, COSDAM (in consultation with the Assessment Development Committee) shall solicit input from technical and subject matter experts to determine whether changes to the content ALDs are warranted or whether a new standard setting shall be conducted, making clear the potential risk of changing cut scores to trends and assessment of educational progress. Relevant factors may include but not be limited to: substantive changes in the item types or in the balance of item types; changes in the mode of administering assessments; advances in standard setting methodologies; and changes in the policy environment for using NAEP results.
b) Within the period for a review of achievement level descriptions and cut scores, changes may occur to a NAEP framework. If a framework is replaced or revised for a major update, a new achievement level setting process may be implemented, except in circumstances where scale score trends are maintained. In this latter instance, COSDAM shall determine how to revise the ALDs and review the cut scores to ensure that they remain reasonable and meaningful.

c) If there are major updates to a NAEP framework, the ALDs shall be updated by the Framework Visioning and Development Panel. (See the Governing Board Policy on Framework Development for additional details). Following an assessment administration under the revised framework, COSDAM shall use empirical data to revise content ALDs to align with the revised framework.

d) As additional validation evidence becomes available, the Board shall review it and make a determination about whether the achievement levels should be reviewed and potentially revised.

Principle 5: Stakeholder Input

The process of developing student achievement levels is a widely inclusive activity. The Governing Board shall provide opportunities to engage multiple stakeholders throughout the achievement level setting process and shall strive to maximize transparency of the process.

a) The process of seeking nominations for the achievement level setting panels shall include outreach to relevant constituencies, such as: state and local educators; curriculum specialists; business representatives; and professional associations in a given content area.

b) The Design Document (describing in detail all planned procedures for the project) shall be distributed for review by a broad constituency and shall be disseminated in sufficient time to allow for a thoughtful response from those who wish to provide one. All interested stakeholders shall have an opportunity to provide public comment.

c) Achievement level setting panelists shall include teachers, non-teacher educators, and other interested members of the general public with relevant educational background and experience, including parents, researchers, and employers. Each panel shall reflect diversity in terms of gender, race/ethnicity, region of the country, urbanicity, and experience with students with disabilities and English language learners.

d) All achievement level setting activities shall be informed by technical advice throughout the process. The Technical Advisory Committee on Standard Setting shall provide ongoing technical input from standard setting and assessment experts, and other groups with relevant technical expertise may be consulted periodically as needed.
e) Ongoing input and coordination with staff and contractors from the National Center for Education Statistics (NCES) is necessary to ensure that all achievement level setting activities are carried out in a manner that is consistent with the design, analysis, and reporting of NAEP assessments.

**Principle 6: Role of the Governing Board**

The Governing Board, through its Committee on Standards, Design and Methodology (COSDAM), shall monitor the development and review of student achievement levels to ensure that the final achievement level descriptions, cut scores, and exemplars recommended to the Governing Board for adoption comply with this policy.

a) The Committee on Standards, Design and Methodology (COSDAM) shall be responsible for monitoring the development and review of achievement levels that result in recommendations to the Governing Board for any NAEP assessment under consideration. COSDAM shall provide direction to the achievement level setting contractor, via Governing Board staff. This guidance shall ensure compliance with the NAEP legislation, Governing Board policies, Department of Education and government-wide regulations, and requirements of the contract(s) used to implement the achievement level setting project.

b) If there is a need to revise the initial achievement level descriptions (ALDs) created at the time of framework development for use in achievement level setting and/or reporting, the Governing Board shall take final action on revised ALDs based on recommendations from COSDAM.

c) COSDAM shall receive regular reports on the progress of achievement level setting projects.

d) COSDAM shall review and formally approve the Design Document that describes all planned procedures for an achievement level setting project.

e) At the conclusion of the achievement level setting project, the Governing Board shall take final action on the recommended cut scores, exemplars, and ALDs for use in reporting. The Governing Board shall make the final determination on the NAEP achievement levels. In addition to the panel recommendations, the Board may consider other pertinent information to assess reasonableness of the results, such as comparisons to other relevant assessments.

f) Following adoption by the Governing Board, the final ALDs, cut scores, and exemplars shall be provided to the National Center for Education Statistics (NCES) for reporting the results of the NAEP assessment(s) under consideration.

g) Consistent with Principle 4 above, COSDAM shall periodically review existing achievement levels to determine whether it is necessary to revise achievement level descriptions or conduct a new standard setting.
National Assessment Governing Board

Developing Student Achievement Levels for the National Assessment of Educational Progress

Policy Statement

It is the policy of the National Assessment Governing Board to conduct a comprehensive, inclusive, and deliberative process to develop and review student achievement levels for the National Assessment of Educational Progress (NAEP).\(^1\) Achievement levels consist of general policy definitions for the NAEP Basic, NAEP Proficient, and NAEP Advanced levels, specific achievement level descriptions (ALDs) for each assessment, cut scores that demarcate adjacent levels, and exemplar items or tasks that illustrate performance at each level. This process shall be conducted according to widely accepted professional standards, to produce results that are reasonable, useful, and informative to the public.

The Governing Board, through its Committee on Standards, Design and Methodology (COSDAM), shall monitor the development and review of student achievement levels to ensure that the final Governing Board-adopted achievement level descriptions, cut scores, and exemplars comply with all principles of this policy.

The achievement level setting process shall be carried out by contractors selected through a competitive bidding process. The process shall be managed in a technically sound, efficient, cost-effective manner, and shall be completed in a timely fashion.

Introduction

Since its creation by Congress in 1988, the Governing Board has been responsible for developing student achievement levels for NAEP assessments. The Governing Board has carried out this important statutory responsibility by engaging with a broad spectrum of stakeholders to develop student achievement levels.

\(^1\) According to current NAEP legislation, the Governing Board shall develop achievement levels for all NAEP assessments except for the Long-Term Trend assessment.
Under provisions of the National Assessment of Educational Progress Authorization Act of 2002 (P.L. 107-279), Congress authorized the Governing Board to, develop, “achievement levels that are consistent with relevant widely accepted professional assessment standards and based on the appropriate level of subject matter knowledge for grade levels to be assessed” (Section 303(e)(2)(A)(i)(I)).

Given this mandate, the Governing Board must ensure that all achievement level setting processes align with current best practices in standard setting, and that appropriate validity evidence is collected and documented to support the intended uses and interpretations of NAEP achievement levels.

The Governing Board has established the following policy definitions for the NAEP achievement levels, as expectations of what students should know and be able to do. They shall be consistent across all assessments in which achievement levels are set.

**NAEP Basic**

This level denotes partial mastery of prerequisite knowledge and skills that are fundamental for performance at the NAEP Proficient level.

**NAEP Proficient**

This level represents solid academic performance for each NAEP assessment. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real world situations, and analytical skills appropriate to the subject matter.

**NAEP Advanced**

This level signifies superior performance beyond NAEP Proficient.

The Governing Board engages multiple stakeholders throughout the achievement level setting process, including:

- Teachers
- Curriculum Experts
- Content Experts
- Assessment Specialists
- State Administrators
- Local School Administrators
- Policymakers
- Business Representatives
- Parents
- Users of Assessment Data
- Researchers and Technical Experts
- Members of the Public

This policy also complies with the documents listed below which express widely accepted technical and professional standards for achievement level setting. These standards reflect the agreement of recognized experts in the field, as well as the policy positions of major professional and technical associations concerned with educational testing.
In conjunction with this policy the Board shall maintain a procedures manual to establish and document additional details about how this policy is to be implemented. As professional standards evolve and new consensus documents are released, this policy and the procedures manual shall be updated to the extent that new professional standards require. Resources for this purpose shall include, but not be limited to the following:


Principles for Setting Achievement Levels

Principle 1: Elements of Achievement Levels

Principle 2: Development of Achievement Level Recommendations

Principle 3: Validation and Reporting of Achievement Level Results

Principle 4: Periodic Review of Achievement Levels

Principle 5: Stakeholder Input

Principle 6: Role of the Governing Board
Principle 1: Elements of Achievement Levels

The Governing Board is responsible for developing student achievement levels for each NAEP assessment. Achievement levels for each NAEP assessment consist of content achievement level descriptions (ALDs), cut scores that demarcate adjacent levels, and exemplar items or tasks that illustrate performance at each level.

a) **Content achievement level descriptions (ALDs)** translate the policy definitions into specific expectations about student knowledge and skills in a particular content area, at each achievement level, for each subject and grade. Content ALDs provide descriptions of specific expected knowledge, skills, or abilities of students performing at each achievement level. Content ALDs reflect the range of performance that items and tasks should measure. During the achievement level setting process, the purpose of content ALDs is to provide consistency and specificity for panelist interpretations of policy definitions for a given assessment. During reporting, content ALDs communicate the specific knowledge and skills represented by *NAEP Basic*, *NAEP Proficient*, and *NAEP Advanced* for a given assessment.

b) **Cut scores** mark the minimum threshold score, the lower bound, for each achievement level. Performance within a given achievement level begins at the cut score for that level and ends just below the cut score for the successive achievement level.

c) **Exemplar items or tasks**, including student responses, illustrate student performance within each of the achievement levels. They provide specific examples to help the public better understand what students in each achievement level know and can do.

Principle 2: Development of Achievement Level Recommendations

The Governing Board shall develop student achievement levels for NAEP, consistent with relevant widely accepted professional assessment standards, based on the appropriate level of subject matter knowledge.

a) A **Design Document** shall be developed at the beginning of the achievement level setting process, to describe in detail the scope of the achievement level setting project being undertaken, including but not limited to all planned materials, procedures, and analyses needed for the project. The Design Document shall be posted for public review with sufficient time to allow for a response from those who wish to provide one.

b) The development of **content achievement level descriptions (ALDs)** shall be completed initially through the process that develops the assessment frameworks. (See the Governing Board Policy on Framework Development for additional details). The Board may then review and revise content ALDs to advance the purposes they serve, whether that is guiding an achievement level setting or informing the public about the meaning of achievement levels. Whether revised or not, the ALDs that guide
achievement level setting shall be articulated in terms of what students should know and be able to do. There shall be no content ALDs developed for performance below the NAEP Basic level.

c) An achievement-level setting panel of subject matter experts shall be convened to recommend achievement level cut scores and exemplars.

i. Each panel shall reflect diversity in terms of gender, race/ethnicity, region of the country, urbanicity, and experience with students with disabilities and English language learners. To ensure that they are qualified to make the judgments required by the achievement level setting process, individual panel members shall have expertise and experience in the specific content area in which the levels are being developed, expertise and experience in the education of students at the grade under consideration, and a general knowledge of assessment, curriculum, and student performance.

ii. This panel shall include teachers, non-teacher educators, and other interested members of the general public with relevant educational background and experience. Teachers shall comprise the majority of the panel, with non-teacher educators (e.g., curriculum directors, academic coaches, principals) accounting for no more than half the number of teachers. The remaining panelists shall be non-educators who represent the perspectives of additional stakeholders representing the general public, including parents, researchers, and employers.

iii. The size of the panels shall reflect best practice in standard setting and be operationally feasible while being large enough to allow for split panels. Most NAEP achievement level settings have historically included approximately 20-30 panelists per grade, divided into two comparable groups with a subset of shared items.

d) Panelists shall receive training on all aspects of the achievement levels setting process to ensure that panelists are well-prepared to perform the achievement level setting tasks required of them. Panelists shall be instructed that their role is to make achievement level recommendations to the Governing Board. Training shall include but not be limited to: the purpose and significance of setting achievement levels for NAEP; the NAEP assessment framework for the given subject area; and administration of a sample assessment under NAEP-like conditions that students experience. It is important for panelists to arrive at a common conceptualization of NAEP Basic, NAEP Proficient, and NAEP Advanced based on the content ALDs. Panelists shall be trained on each element of the judgmental task they perform, including the selection of exemplar items. They should be led by capable content facilitators (who are content experts and have previous experience with achievement level setting) and process facilitators (who have background in standard setting and experience leading panelists through the achievement level setting process). Facilitators shall take a neutral stance and not attempt to influence panelist judgments.

Comment [RS4]: There was a suggestion to delete this phrase or provide more information by saying that the item pool is also divided into two equivalent sets with a subset evaluated by both panel groups. The latter is information that seems more appropriate for a procedures manual.
e) The achievement level setting method that generates cut score recommendations may differ depending upon the specific assessment. Nevertheless, the method shall have a solid research base and be appropriate for the content area, item types, number of items, scoring rubrics, and mode of administration, as applicable.

f) Evaluations shall be administered to panelists throughout the achievement level setting process, in accordance with current best practices. Evaluations shall be part of every major component of the process, and panelists shall be asked to confirm their readiness for performing their tasks. Evaluation data may be used for formative purposes (to improve training and procedures in future meetings); summative purposes (to evaluate how well the process was conducted and provide procedural validity evidence); and to inform the Governing Board of any relevant information that could be useful when considering cut score recommendations. The panelists shall have an opportunity to indicate to the Board whether they believe the recommended cut scores are reasonable.

g) In accordance with current best practices, feedback shall be provided to panelists, including “impact data” (i.e., the implications of their selected cut scores on the reported percentages of students at or above each achievement level).

h) The process shall consist of at least two achievement level setting meetings with distinct groups of panelists, a pilot study, and an operational meeting. The purpose of the pilot study is to conduct a full “dress rehearsal” of the operational meeting, including but not limited to: an opportunity to test-try out materials, training procedures, collection of panelist judgments, feedback given to panelists through the process, software used to conduct analyses, meeting logistics, and other essential elements of the process. The pilot study may result in minor changes to the procedures, as well as major changes that would need additional study before being implemented in an operational meeting. The pilot study provides an opportunity for procedural validity evidence and to improve the operational meeting. At the discretion of the Governing Board, other smaller-scale studies may be conducted prior to the pilot study or in response to issues raised by the pilot study. The criteria in Principle 2a apply to panelists of both meetings.

i) The Governing Board shall ensure that a Technical Advisory Committee on Standard Setting (TACSS) is convened to provide technical advice on all achievement level setting activities. Technical advice provided by standard setting experts throughout the project is intended to ensure that all procedures, materials, and reports are carried out in accordance with current best practices, providing additional validity evidence for the process and results. The Board or its contractor may also seek technical advice from other groups as appropriate, including NCES and the larger measurement community (e.g., the National Council on Measurement in Education).

j) All aspects of the procedures shall have documentation as evidence of the appropriateness of the procedures and results. This evidence shall be made available to the Board by the time of deliberations about the achievement levels. A summary of the evidence shall be available to the public when the achievement level results are reported.
k) Sample items and student responses known as exemplars shall be chosen from the pool of released items for the current NAEP assessment to reflect performance in the NAEP Basic, NAEP Proficient, and NAEP Advanced regions of the scale. The use of exemplars is intended to help the public better understand what students who are in performance at an achievement level actually know and are able to demonstrate for each subject and grade. When possible, exemplars may also be chosen that reflect performance at threshold scores. The collection of exemplars shall reflect the content found in the achievement level descriptions and the range of item formats on the assessment.

l) The outcomes from the achievement level setting panel meetings (recommended cut scores, exemplars, and ALDs for use in reporting) shall be forwarded to the Board for their consideration.

Principle 3: Validation and Reporting of Achievement Level Results

The achievement level setting process shall produce results that have validity evidence for the intended uses and interpretations and are informative to policy makers, educators, and the public.

a) Professional testing standards require evidence to support the intended interpretations and uses of test scores. Among the sources of evidence supporting the validity of test scores is evidence bearing on the standard setting process and results. Although standard setting is necessarily judgmental, and with no “true” or “correct” cut scores, the Board shall examine and consider available evidence about the procedural integrity of the achievement level setting process, the reasonableness of results, and other evidence in order to support intended uses and interpretations.

b) The Board shall examine and consider all evidence related to validity of the achievement level setting activities. These data shall include, but not be limited to: procedural evidence such as training, materials and panelist evaluation data; reliability evidence such as consistency across panelist type, subpanels, rounds, and meetings, if appropriate; and external comparisons to other similar assessments, if appropriate, with necessary caveats. The results from validation efforts shall be made available to the Board in a timely manner so that the Board has access to as much validation data as possible as it considers the recommendations regarding the final levels.

b) NAEP achievement levels are intended to estimate the percentage of students (overall and for selected student groups) in each achievement level category, for the nation, and for states and trial urban districts (TUDAs) for some assessments. NAEP is prohibited by law from reporting any results for individual students or schools, so achievement levels do not apply to individual students or schools.

c) To facilitate valid uses of ALDs for reporting, the Board shall ensure that the
descriptions of performance for the achievement levels reflect what the empirical data reveal about the knowledge and skills of students in that score range. The Board shall revisit and may revise content ALDs following the achievement level setting to ensure that they are consistent with empirical evidence of student performance. In particular, when content ALDs are reported with results, they shall be written to incorporate empirical data from student performance. They shall describe what students at each level do know and can do rather than what they should know and should be able to do.

e) The Board shall examine and consider all evidence related to validity of the achievement level setting activities. These data shall include, but not be limited to: procedural evidence such as training, materials and panelist evaluation data; reliability evidence such as consistency across panelist type, subpanels, rounds, and meetings, if appropriate; and external comparisons to other similar assessments, if appropriate, with necessary caveats. The results from validation efforts shall be made available to the Board in a timely manner so that the Board has access to as much validation data as possible as it considers the recommendations regarding the final levels.

f) In describing student performance using the achievement levels, terms such as “students performing at the NAEP Basic level” or “students performing at the NAEP Proficient level” are preferred over “Basic students” or “Proficient students”. The former implies that students have mastery of particular content represented by the achievement levels, while the latter implies an inherent characteristic of individual students.

g) In reporting the results of NAEP, the three achievement levels of NAEP Basic, NAEP Proficient, and NAEP Advanced refer to the three regions of the NAEP scale at and above each respective cut score. The remaining region that falls below the NAEP Basic cut score shall be identified as “below NAEP Basic” when a descriptor is necessary.

f) In describing the NAEP Proficient level, reports shall emphasize that the policy definition is not intended to reflect “grade level” performance expectations, which are typically defined normatively and can vary widely by state and over time. NAEP Proficient may convey a different meaning from other uses of the term “proficient” in common terminology or in reference to other assessments.

g) To facilitate valid uses of ALDs for the purpose of reporting, the Board shall ensure that the descriptions of performance for the achievement levels reflect what the empirical data reveal about the knowledge and skills demonstrated by students in that score range. To develop ALDs for reporting, following the achievement level setting the Board shall revisit and may revise content ALDs to ensure that they are consistent with empirical evidence of student performance. In particular, these “Reporting ALDs” chosen to illustrate the knowledge and skills demonstrated at different achievement levels shall be written to incorporate empirical data from student performance. Reporting ALDs shall describe what students at each level do know and can do rather than what they should know and should be able to do.
h) An interpretative guide shall accompany NAEP reports, including specific examples of appropriate and inappropriate interpretations and uses of the results.

**Principle 4: Periodic Review of Achievement Levels**

Periodic reviews of existing achievement levels shall determine whether new achievement level descriptions and/or cut scores are needed to continue valid and reliable measurement of current student performance and trends over time.

a) At least once every 10 years or 3 administrations of an assessment, whichever comes later, the Governing Board, through its Committee on Standards, Design and Methodology (COSDAM), shall review the alignment between the content ALDs and items, based on empirical data from recent administrations of NAEP assessments. In its review, COSDAM (in consultation with the Assessment Development Committee) shall solicit input from technical and subject matter experts to determine whether changes to the content ALDs are warranted or whether a new standard setting shall be conducted, making clear the potential risk of changing cut scores to trends and assessment of educational progress. Relevant factors may include but not be limited to: substantive changes in the item types or in the balance of item types; changes in the mode of administering assessments; advances in standard setting methodologies; and changes in the policy environment for using NAEP results.

b) Within the period for a review of achievement level descriptions and cut scores, changes may occur to a NAEP framework. If a framework is replaced or revised for a major update, a new achievement level setting process may be implemented, except in circumstances where scale score trends are maintained. In this latter instance, COSDAM shall determine how to revise the ALDs and review the cut scores to ensure that they remain reasonable and meaningful.

c) If there are major updates to a NAEP framework, the ALDs shall be updated by the Framework Visioning and Development Panel. (See the Governing Board Policy on Framework Development for additional details). Following an assessment administration under the revised framework, COSDAM shall use empirical data to revise content ALDs to align with the revised framework.

d) As additional validation evidence becomes available, the Board shall review it and make a determination about whether the achievement levels should be reviewed and potentially revised.

**Principle 5: Stakeholder Input**

The process of developing student achievement levels is a widely inclusive activity. The Governing Board shall provide opportunities to engage multiple stakeholders throughout the achievement level setting process and shall strive to
maximize transparency of the process.

a) The process of seeking nominations for the achievement level setting panels shall include outreach to relevant constituencies, such as: state and local educators; curriculum specialists; business representatives; and professional associations in a given content area.

b) The Design Document (describing in detail all planned procedures for the project) shall be distributed for review by a broad constituency and shall be disseminated in sufficient time to allow for a thoughtful response from those who wish to provide one. All interested stakeholders shall have an opportunity to provide public comment.

c) Achievement level setting panelists shall include teachers, non-teacher educators, and other interested members of the general public with relevant educational background and experience, including parents, researchers, and employers. Each panel shall reflect diversity in terms of gender, race/ethnicity, region of the country, urbanicity, and experience with students with disabilities and English language learners.

d) All achievement level setting activities shall be informed by technical advice throughout the process. The Technical Advisory Committee on Standard Setting shall provide ongoing technical input from standard setting and assessment experts, and other groups with relevant technical expertise may be consulted periodically as needed.

e) Ongoing input and coordination with staff and contractors from the National Center for Education Statistics (NCES) is necessary to ensure that all achievement level setting activities are carried out in a manner that is consistent with the design, analysis, and reporting of NAEP assessments.

f) The Governing Board may ask its standing groups representing various constituencies to provide input on the achievement level setting process.

Principle 6: Role of the Governing Board

The Governing Board, through its Committee on Standards, Design and Methodology (COSDAM), shall monitor the development and review of student achievement levels to ensure that the final achievement level descriptions, cut scores, and exemplars recommended to the Governing Board for adoption comply with this policy.

a) The Committee on Standards, Design and Methodology (COSDAM) shall be responsible for monitoring the development and review of achievement levels that result in recommendations to the Governing Board for any NAEP assessment under consideration. COSDAM shall provide direction to the achievement level setting contractor, via Governing Board staff. This guidance shall ensure compliance with the NAEP legislation, Governing Board policies, Department of Education and government-wide regulations, and requirements of the contract(s)
used to implement the achievement level setting project.

b) If there is a need to revise the initial achievement level descriptions (ALDs) created at the time of framework development for use in achievement level setting and/or reporting, the Governing Board shall take final action on revised ALDs based on recommendations from COSDAM.

c) COSDAM shall receive regular reports on the progress of achievement level setting projects.

d) COSDAM shall review and formally approve the Design Document that describes all planned procedures for an achievement level setting project.

e) At the conclusion of the achievement level setting project, the Governing Board shall take final action on the recommended cut scores, exemplars, and ALDs for use in reporting. The Governing Board shall make the final determination on the NAEP achievement levels. In addition to the panel recommendations, the Board may consider other pertinent information to assess reasonableness of the results, such as comparisons to other similar relevant assessments.

f) Following adoption by the Governing Board, the final ALDs, cut scores, and exemplars shall be provided to the National Center for Education Statistics (NCES) for reporting the results of the NAEP assessment(s) under consideration.

g) Consistent with Principle 4 above, COSDAM shall periodically review existing achievement levels to determine whether it is necessary to revise achievement level descriptions or conduct a new standard setting.
September 29, 2018

National Assessment Governing Board
U.S. Department of Education
800 North Capitol Street NW – Suite 825
Washington, DC 20002-4233
Attention: Sharyn Rosenberg, Assistant Director

Comments on Revised Policy on Developing Student Achievement Levels for NAEP
Document Citation: 83 FR 45618

The Council of the Great City Schools (Council), the coalition of the nation’s largest central city school districts, submits the following comments on the revised policy on Developing Student Achievement Levels for the National Assessment of Educational Progress (NAEP) in response to the September 10, 2018 notice in the Federal Register. Over the years, the Council has worked closely with the National Assessment Governing Board (Governing Board) and the National Center for Education Statistics (NCES) on a variety of efforts to measure and improve student learning outcomes. Therefore, the Council supports the Governing Board’s efforts to periodically review and update NAEP policies and practices.

The Council is dedicated to the improvement of education for children in the nation’s inner cities. The Council and its member districts work to help our public-school children meet the highest standards and become successful and productive members of society. The organization and the 27 Trial Urban District Assessment (TUDA) participants regularly use results from NAEP to measure our progress and relative standing in achieving our goals. In fact, the Council of the Great City Schools initiated TUDA in 2000 as a way of holding ourselves and our students to the highest standards. As a result, we are heavily invested in any changes in policies and practices related to NAEP and other national measures of educational progress.

The Council’s comments in this letter are focused on retaining a rigorous assessment, maintaining current terminology in NAEP achievement levels, thoughtfully tagging NAEP performance levels to meaningful high-level expectations, and devoting additional effort to explaining to the public what NAEP is intended to measure and what it is not. Please let us know if clarification is needed on any of these comments.

Sincerely,

Michael Casserly
Executive Director
COUNCIL OF THE GREAT CITY SCHOOLS COMMENTS
ON THE REVISED POLICY ON DEVELOPING STUDENT
ACHIEVEMENT LEVELS FOR THE
NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

New Principle on Periodic Review of Achievement Levels and Cut Scores

The Council of the Great City Schools agrees that the National Assessment Governing Board (NAGB) and the National Center for Educational Statistics (NCES) should periodically review the performance levels used in the National Assessment of Educational Progress (NAEP). These periodic reviews, however, should be tempered by the continuing need to have a national assessment that is consistently applied from state-to-state, a national assessment that allows comparisons across states and participating TUDA participants, and a national assessment that provides a clear trend line over time for the nation, states, and districts.

That said, the Council is also interested in seeing NAGB and NCES devote additional time and effort to benchmarking NAEP performance levels to such concrete high-level expectations as current college- and career-readiness standards, international measures of performance, average entrance requirements to competitive colleges, or other publicly understandable measures of excellence, rather than the judgment of expert panels, even if the result is somewhat aspirational. We believe the standard against which NAEP is pegged ought to be rigorous and not reflect the lowest common denominator as some have argued. Above all, NAEP is a measure of how the nation, its states, and many of its critical large city school systems are performing educationally; it is not an accountability tool or an assessment of individual attainment. To that end, NAEP should be pegged to the highest possible but specific standard of attainment.

We also encourage the Governing Board NOT to adjust cut scores or descriptions without extensive research or overwhelmingly compelling evidence to suggest misalignment with desired interpretations of NAEP results. The Council and our member districts that participate in the TUDA program find the current achievement level cut scores for proficiency to be a fair and accurate expectation of student outcomes. We believe our schools, teachers, and administrators should have the highest expectations for student achievement, and we believe the current level of proficiency reflects a realistic, although high, expectation for our students. Students are harmed when expectations are low, not high, a situation that urban schools are more aware of than many others.

While additional, more robust, research is needed on the current proficiency levels, a cursory review of post-secondary outcomes suggests current proficiency standards are basically sound. According to the 2017 Current Population Survey from the U.S. Census Bureau, 1 about 37 percent of adults 25 to 34 years old had a bachelor’s degree or higher. On the 2007 eighth grade NAEP assessment, about the same time those 25-year olds would have been in eighth grade, approximately 32 percent the nation’s students were proficient in reading and math. 2 By design or not, NAEP appears on its face to have some grounding in real-world attainment.

2 Source: U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics.
Arguments in favor of shifting NAEP achievement level cut scores masks the real problem in educational attainment across the country. In 2017, over 57 percent of students not eligible for a free or reduced-price lunch were proficient or better on the fourth grade NAEP mathematics assessment compared to 24.7 percent of students eligible for a free or reduced-price lunch. The disparities are even more discouraging when race and poverty are considered jointly. Advocates calling for the lowering of proficiency standards shift the nation’s attention away from the real issues about disparities in educational outcomes among the nation’s impoverished and traditionally under-represented student groups in favor of making more students looking artificially higher performing.

Finally, the Council does wonder whether there is a way to retain the current performance levels but expand use of NAEP’s 500-point scale. At present, scores routinely fall between points 200 and 300 on the scale, making it difficult to show movement—either upwards or downwards. A robust discussion about the wisdom of this seems prudent in the current review.

**Recommendation:** The Council recommends retaining NAEP’s high level of rigor in defining proficiency, but it also proposes that NAGB and NCES conduct additional, robust research to better tie those proficiency levels to college- and career-ready standards, some international benchmark, or to post-secondary success.

**Change in Terminology from Proficient to NAEP Proficient**

The Council also does not think that the Governing Board should change the terminology used by NAEP from Proficient to NAEP Proficient, from Basic to NAEP Basic, or from Advanced to NAEP Advanced. We think the current labels should remain for several reasons. First, the summary of proposed revisions suggests that NAEP achievement levels should be “better differentiated” from other common uses of the terms Basic, Proficient, and Advanced. For decades, however, NAEP has been, and should remain, the standard for these terms. Application of these terms from assessment-to-assessment have been made relative to NAEP definitions—even if they have not been faithfully applied. Changing the terminology suggests that NAEP should no longer be the standard upon which we understand student achievement.

Second, introducing NAEP-specific Basic, Proficient, and Advanced levels might introduce considerable confusion to the public’s understanding of student achievement across assessments. In a review of NAEP Achievement Levels conducted by the National Academies of Science, Engineering, and Medicine, they conclude that, “during their 24 years [the achievement levels] have acquired meaning for NAEP’s various audiences and stakeholders; they serve as stable benchmarks for monitoring achievement trends, and they are widely used to inform public discourse and policy decisions. Users regard them as a regular, permanent feature of the NAEP reports” (Edley & Koenig, 2016; p. Sum-8). The public’s understanding of current terminology is well entrenched and already commonly understood. Parents, educators, and the public are better served when the educational community can consistently articulate student achievement outcomes. The Council encourages NAGB to retain current terminology without the modifying

---

terminology, because NAEP should remain the standard against which these proficiency levels should be defined.

Recommendation: In 83 FR 45618 strike all references to “NAEP Basic, NAEP Proficient, and NAEP Advanced” in the document and restore the use of the terms Basic, Proficient, and Advanced.

Communicating NAEP Terminology

While there is long-standing public and professional understanding of NAEP levels, there are also forces who would mislead the public about NAEP either in pursuit of their own agendas or because they were misinformed. There is little way for NAGB or NCES to prevent the deliberate misuse of NAEP results, but both organizations and their partners, including the Council, could do a better job of informing the public about what these performance levels mean and what they don’t mean. For instance, one routinely hears that NAEP levels are akin to grade-level scores. But, in the announcement of results every two years, there is little time devoted to revisiting or describing the definitions of the performance levels or how they were arrived at. Having the performance levels tagged to some external benchmark, as we suggested in the earlier recommendation, might help NAGB and NCES better describe what NAEP means. And spending some time during the release on what the performance levels mean—beyond examples of what students can do under each level—might help ward off some misunderstanding of the levels and protect against the deliberate misuse of terms.

Recommendation: Devote more time and attention during the release of NAEP results to what the proficiency levels mean and what they don’t mean beyond the examples that are often presented to illustrate performance.
Comments on Revised Policy on Developing Student Achievement Levels for NAEP
Document Citation: 83 FR 45618

The Council of Chief State School Officers (CCSSO) submits the following comments on the revised policy on Developing Student Achievement Levels for the National Assessment of Educational Progress (NAEP) in response to the September 10, 2018 notice in the Federal Register. CCSSO is a nonpartisan, nationwide, nonprofit organization of public officials who head departments of elementary and secondary education in the states, the District of Columbia, the Department of Defense Education Activity, the Bureau of Indian Education and the five U.S. extra-state jurisdictions. CCSSO has a positive and longstanding relationship with the National Assessment Governing Board (NAGB) and the National Center for Education Statistics (NCES).

State leaders recognize the importance of the National Assessment of Educational Progress (NAEP) as the longstanding common assessment metric used across the country. As a membership organization representing state leaders, CCSSO also greatly values this measure as it continues to serve as an important way for our organization to better understand and analyze student academic progress over time at the national level.

For these reasons, CCSSO is supportive of the revised principles as outlined in the notice published in the Federal Register. In particular, we believe the following proposed principles would be helpful to states, districts, and schools.

- Principle 1 (Elements of Achievement Levels) is clear and provides helpful descriptions of content level achievement level descriptions (ALDs), cut scores, and exemplar items and tasks. All of these, taken together, help to provide context and explain the meaning of the achievement levels used by NAEP.

- Principle 2 (Development of Achievement Level Recommendations) provides additional context about the NAEP achievement levels. We support the development of a design document to further explain who should be included on the achievement-level setting panels, particularly in regard to inclusion of diverse individuals in terms of gender, race/ethnicity, region of the country, urbanicity, and experience with students with disabilities and English language learners.

- Principle 4 (Periodic Review of Achievement Levels) commits NAGB to a review of existing achievement levels at least once every 10 years, or three test administrations, whichever comes later. We support this periodic review and believe that this type of examination is healthy for the assessment. We urge caution before proceeding with changes to the current “cut scores” or level of rigor attached to the current achievement levels. While there have been some difficulties in understanding the precise meaning of the current achievement levels, these levels provide an ongoing national trend and accurate state-by-state comparisons. CCSSO would like to ensure that any review maintain the appropriate balance and recognize the importance and value of NAEP as
a longstanding common assessment metric across the country while at the same time providing detailed information to states, districts and schools to use in decision-making.

- Principle 5 (Stakeholder Input) is an important component of NAEP, and we support inclusion of diverse stakeholders for the activities outlined in the principle, including nominations of achievement-level setting panelists, review of design documents, and review of achievement-level setting activities by the Technical Advisory Committee on Standard Setting. We have appreciated the inclusion of state chiefs through the NAGB and CCSSO State Policy Task Force and want to ensure this structure and similar activities continue for state chiefs and other stakeholders in the future.

In general, CCSSO supports the revised policy on Developing Student Achievement Levels for the NAEP as outlined in the notice published in the Federal Register. However, in addition to the principles already outlined, we believe the final policy also should address additional work to examine the alignment between the current NAEP frameworks and the content standards that are used in states. As an organization, CCSSO is dedicated to setting high expectations for all students through college- and career-ready standards. Since 2010, each state has improved their standards.

CCSSO recognizes and appreciates the robust and inclusive procedures by NAGB in support of a meaningful, valid, and reliable administration of the NAEP. At the same time, we support any additional work that can help to continue to explain the meaning and context surrounding NAEP’s achievement levels. For this reason, we urge NAGB to continue to examine the alignment between the NAEP frameworks and the content standards states have adopted since 2010, as we know every state has updated its content standards in NAEP-tested subject areas significantly in recent years.

Mis-alignment presents two challenges to states and to NAEP. First, it is not reasonable for states – as well as districts, schools and teachers – to be held accountable if material that has not yet been taught at a particular grade level. Second, NAEP may not be providing accurate measurement if it fails to test material that has been taught in a particular grade level, but not yet covered by NAEP. We believe examining the alignment of the NAEP frameworks can be done in a way that maintains NAEP’s long-term trend and continues its longstanding value as the ‘Gold Standard’ in assessments across the country, as evidenced in the Evaluation of the Achievement Levels for Mathematics and Reading on the National Assessment of Educational Progress study from 2016. CCSSO stands ready to support and assist throughout this process.

Thank you for your consideration of our comments on the revised policy on Developing Student Achievement Levels for the NAEP in response to the September 10, 2018 notice in the Federal Register.

Sincerely,

Carissa Moffat Miller
Executive Director
Council of Chief State School Officers
October 15, 2018

NAEP Achievement Level Setting Policy
National Assessment Governing Board
800 North Capitol Street, N.W., Suite 825
Washington, DC 20002

RE: Revised Policy on Developing Student Achievement Levels for NAEP

Dear Dr. Rosenberg,

The National Assessment of Educational Progress (NAEP) is one of the most valued and recognized measures of educational quality and equity in the United States. Every two years, the assessment provides educators, policymakers, advocates, and the public with critical data on student achievement that are comparable over time and across states. As state standards and assessments continue to evolve, NAEP’s consistency has made it an invaluable, trustworthy yardstick for gauging trends in educational outcomes for all students and, importantly, individual student groups.

NAEP data are a critical tool for equity advocates, who use the results to draw attention to disparities in achievement both within their states and across the nation. NAEP allows advocates and researchers to identify states that are making strong gains for historically underserved students and learn from their experiences. It allows advocates to maintain the push for high expectations for all students and to benchmark their state’s results against top performers. NAEP achievement levels are crucial for understanding and communicating how well our education system is serving students, and their consistency over time is a major part of the assessment’s utility.

For the reasons above, we write in the interest of ensuring that the National Assessment Governing Board’s (NAGB) final Achievement Level Policy maintains the rigor, comparability and transparency that have made NAEP such a valuable resource for education stakeholders. We especially urge NAGB not to lower expectations for what it means to be “Proficient” or “Advanced” arbitrarily. If the Board deems it necessary to revise achievement level thresholds, we strongly recommend that revised definitions be based in the reality of what students need to know and be able to do at each tested grade level to be prepared for success in postsecondary education without the need for remediation by the time they leave high school.

Our education system has a long history of telling students and their families — especially those from underserved communities — that their schools are preparing them for future success. Meanwhile, college remediation rates remain high and employers struggle to find candidates who have the necessary reading and math skills. One of the biggest “value adds” of NAEP has been its consistent high expectations. If the revision of the Achievement Level Policy results in lower expectations for what it means to be “Proficient” or “Advanced” without solid justification for these changes, it could harm students across the country, with the highest risks for students who are already underserved in our
What’s more, in the absence of a clear rationale for the change, sudden increases or decreases in the percent of students who are “Proficient” may undermine faith in the measure.

Given the potential implications for the usefulness and transparency of NAEP data, **NAGB should only redefine NAEP achievement levels if there is strong evidence that the levels do not currently signify what their definitions entail.** Moreover, if NAGB deems it necessary to revise the levels or what they mean, it should **align the new achievement level thresholds and definitions to what students need to know and be able to do to maximize their educational and career options upon leaving high school (i.e., to be prepared for success in college or postsecondary training without the need for remediation).** In 2014, NAGB conducted a series of studies that showed that the 12th grade proficient cut score on NAEP was well aligned with college-readiness criteria. NAGB could use this research to identify appropriate cut scores for proficiency in eighth and fourth grade – levels that would indicate whether students are on track to meet the 12th grade proficiency (i.e., college-preparedness) expectations. This type of alignment would increase the accuracy of NAEP achievement levels, grounding them in the reality of what students need to know and be able to do, while maintaining transparency. If this alignment process results in different cut scores for NAEP achievement levels, the Nation’s Report Card should allow users to obtain percentages of students scoring at each performance level using previous and any new definitions both for the most recent and prior years’ results.

Any changes to NAEP achievement levels will pose a communications challenge, and changes that lack a clear rationale risk could detract from the utility of the results. We hope that if NAGB decides it is necessary to revise achievement levels or their definitions, it will do so in a way that yields more meaningful information about how well prepared students are for postsecondary success without the need for remediation.

Sincerely,

Daria Hall
Vice President for Partnerships and Engagement
The Education Trust
Dear Ms. Carr,

On behalf of the undersigned organizations, representing public school superintendents, district administrators, and educational service agency administrators, we submit this letter in response to the proposed changes by the National Assessment Governing Board (NAGB) to the National Assessment of Educational Progress (NAEP) achievement benchmarks. We are opposed to the changes as proposed and are concerned the changes fail to meet the goal in the proposal’s first paragraph: to produce achievement levels that are “reasonable, useful, and informative to the public.”

We thank you for extending the comment period. While still short, the additional time reflects the reality of what it takes to review and adequately respond to the type of changes proposed in the rule and what it could mean for the nation’s schools, the students they serve, and the communities to which they are accountable and in which they live.

The original achievement levels were developed in a rushed process, and resulted in levels that continue to confuse educators, citizens, and policymakers. The levels have been described as ‘wishful thinking’ more than ‘reasonable’ or ‘common sense’, and the latest research linking NAEP’s benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP’s proficiency bar.

Our organizations support high standards, and this response is not a request to lower the standards of NAEP. Rather, we are requesting that should NAGB move forward with a proposal to revise the definitions of NAEP terms, that the final product be an improvement over the current definitions and ensure that parents and the community can understand what the NAEP standard terms mean and how those levels do or do not relate to other academic scores, including assessments and grade level. We align with the recommended changes submitted by the National Superintendents Roundtable, as they appear in the section starting at the bottom of page six of their letter. Their recommendations revise the proposed policy definitions and represent changes that improve efforts to better communicate information that is reasonable and accessible to educators and the general public.

We thank you for reviewing our submitted comment today and urge you to revise the policy definitions for the standards. We reiterate that the proposed policy definitions for NAEP achievement levels miss the mark and are unresponsive to criticism of the levels. These standards and benchmarks must be revised so as to be expressed in terms that parents and the public can understand. We welcome the opportunity to work with you as this process moves forward. Please direct any questions to Noelle Ellerson Ng (nellerson@aasa.org).

Sincerely,

Daniel A. Domenech
AASA, The School Superintendents Association

Joan Wade
Association of Educational Service Agencies

---

1 Letter from National Superintendents Roundtable (September 18, 2018)
http://files.constantcontact.com/d6ed868c001/b63f4624-49ad-40a1-afd8-660913afe0cd.pdf
October 15, 2018

National Assessment Governing Board
800 North Capitol Street, N.W., Suite 825
Washington, DC 20002

Dear Governing Board Members,

The Alliance for Excellent Education (All4Ed) appreciates the opportunity to provide feedback regarding the National Assessment Governing Board’s proposed revision to its policy on Developing Student Achievement Levels for the National Assessment of Educational Progress (NAEP). NAGB’s policy on setting achievement levels for NAEP was last updated in 1995. As such, it is pertinent to review and revise the policy as necessary. All4Ed is pleased that the current proposal does so in such a way that maintains NAEP’s critical role as an honest gauge of student performance across the nation.

It is critical that NAEP’s achievement levels continue to be rigorous. While the nation has yet to provide the majority of students with sufficient support to reach the levels of NAEP Proficient or NAEP Advanced, this does not mean that these levels should be lowered. The 21st century economy demands that more students acquire greater levels of education than ever before – content mastery, critical thinking, the ability to apply knowledge to solve problems, and other “deeper learning” competencies. These competencies, demonstrated at the levels of NAEP Proficient and NAEP Advanced, should be the goal to which the education system aspires.

Moreover, NAEP’s results are consistent with the results of other nationally administered assessments. This demonstrates that NAEP’s achievement levels do not hold students to an unreasonable standard. When comparing the percentage of students performing at the level of NAEP Proficient or higher with the percentage of students receiving a four or higher on the Partnership for Assessment of Readiness for College and Careers (PARCC) assessment, the results are within a modest margin (roughly 34 to 40 percent). In addition, the percentage of students who are prepared for college according to NAEP, the SAT, and the ACT are also within a modest margin (37 to 46 percent). Thus, NAEP is setting a similar bar to assessments that are widely acknowledged to measure college readiness.

For decades, NAEP has been the standard-bearer for academic performance. The growing diversity of statewide assessments, indicators of college and career readiness, and accountability systems increases the importance of NAEP as a consistent state and national measure to which other assessments can be compared. NAGB’s proposed policy allows NAEP to continue this critical role. Lowering expectations would be a disservice to students and the nation at large.

Sincerely,

Bob Wise
President
"Go West...." Horace Greeley reportedly (arguably) told Josiah Grinnell to "go west" in an instance of some of the most important directional advice ever given. The advice was not "go west 182 degrees for 500 miles and then 187 degrees west" with 95 percent confidence in that direction. And yes, Josiah Grinnell said "It is a wide country, but I do not know just where to go." However, there was an important, compelling rationale for "going west" and the precise direction for going west in this "wide country" was less important and not entirely mapped out. But Greeley's advice was given in the mid-nineteenth century, not in the 1990s when NAEP achievement levels were being developed and when latitude and longitude measures would have been expected.

The comparison to "go west" and to chart your educational direction with a NAEP compass where "Basic" and "Proficient" are the most relevant markers is not a perfect comparison, and yet few comparisons are.

As a participant in the original National Assessment Governing Board process to establish achievement levels, I can say that I was certain beyond a reasonable doubt that creating achievement levels was the right direction for NAEP and the nation. I was less certain that the efforts to describe quite precise "latitude/longitude" achievement level measures could be superimposed on a judgment decision ... a statutorily described judgment decision.

It is important, I would assert vital, to remember where our nation was in educational measurement when NAEP achievement levels were produced. In 1984 I reported to the Southern Regional Education Board and its chair, Governor Lamar Alexander, that virtually all of the southern states had student achievement scores "at or above the national average" on the most widely recognized and used "national" education tests. Yes, southern states were among the nation's poorest by most economic measures and were usually cited among the lowest in various educational measures, and almost all of these states were reporting to their residents (including parents of course) that the state's student test results were "at or above the national average" in some subjects and at some grades.

I never claimed that this was a conspiracy by state and district education leaders. A West Virginia physician did claim that it was an unethical conspiracy a year or so later with a report asserting that every state was reporting that its students' achievement was "at or above the national average". So in the 1980s we had state reports showing student achievement in virtually every state "at or above the national average" (not to mention that there were "national norms" and "urban national norms") and the U.S. Department of Education's widely publicized "Wall Chart" with a very different and equally misleading picture of American education. Then add to this mix the numerous state "minimum competency tests" that had virtually by definition low standards and high percentages of students meeting these "competency" standards.

There was certainly plenty of legitimate criticism of the "Wall Chart" and the misleading state SAT scores it featured that created bizarre state rankings. There were also legitimate criticism of the minimum competency tests, but often the criticism was more about the high rate of
failure for economically disadvantaged students. This was not only a concern about testing, but I would argue a major, fundamental issue for our nation and our educational system. What "passing" state minimum competency tests meant, and didn't mean, for students and the nation was also a significant problem.

Shortly after the first NAEP reports based on achievement levels were available I published an SREB report "Setting Education Standards High Enough" (1996). At that time there were eleven states that I could identify with their own eighth grade mathematics standards that were NOT described as "minimum" expectation standards. The percentages of students meeting their own state's standard ranged from 13 percent to 84 percent. Therefore, in one state residents (including parents) were being told that 87 percent of their students DID NOT MEET the standard and in another state residents (including parents) were being told that 84 percent of their students DID MEET their state standard.

If this is confusing, add that when the first eighth grade NAEP achievement level results were available in the mid-1990s it was possible to look at the results for the state where 13 percent of students met its standard (Delaware) and a state where 83 percent met its standard (Georgia). And what did one find? The eighth graders in Delaware had higher NAEP mathematics scores than did Georgia eighth graders.

This was the situation -- I would argue very confusing situation for the American public (and parents)-- in which the NAEP achievement levels were created.

If a judgment decision (achievement levels) about a measurement device (NAEP) is to be seen as strictly or overwhelmingly a psychometric matter then the achievement levels may never be deemed "correct". My view, in the 1990s when helping establish the achievement levels... and now... is that this process is more about direction than precision and the results should be used in that way. I never argued that the achievement levels were the defining measures for states. I did, and do, believe that state education and government leaders should continually seek and analyze state-NAEP results to determine if their "official" student assessment results and their state-NAEP results are pointing in the same direction. If so, what does this mean? If not, what does this mean? If there is a gap in these results, what does this mean?

Yes, state-NAEP results can be misused and state leaders and the press should strongly oppose this. This possible misuse is no greater than the obvious misuse of non-NAEP results before the 1990s, and no student has ever failed (or passed) a NAEP test and had to deal with the consequences. I am not denying the possibility of misuse, intentional or unintentional but NAEP achievement levels did not create the misuse of test results.

My greatest concern about the possible misuse of NAEP achievement level results is the one about which I have the most serious disagreement with some achievement level critics. Some critics claimed, and from the most recent report apparently still do, that the achievement levels were created to discredit public education by producing very low measures of student achievement. When National Association Governing Board members totally committed to public education said "No" to this charge, critics claimed that the Board was duped by one or
two persons. None of these critics were at all of the Board's discussions of achievement levels. The Governing Board had some "super smart" folks and even if there were some "super dupers" they could not have, and did not, hoodwink the Governing Board. Some persons might have sought to misuse the results but it was not the Governing Board's intent.

For NAEP and its achievement levels, "precision" measures are a legitimate touchstone, but it is a judgment process with direction, stability...and transparency...being vitally important factors that are not given to precise measures. "Transparency" means showing the American public (and of course parents) NAEP tests and many NAEP test questions/answers so that over time there are ways they can participate in the "judgment decision" about NAEP achievement levels. This NAEP transparency and sharing, not equating studies with other nations results, is the most important missing input about NAEP achievement levels.

Mark Musick
former member and chair
National Assessment Governing Board
October 11, 2018

Sharyn Rosenberg
NAEP Achievement Level Setting Policy
National Assessment Governing Board
800 North Capitol Street NW, Suite 825
Washington, DC 20002

Document Number: 2018-21451

Dear Ms. Rosenberg:

Thank you for the opportunity to comment on the revised policy on Developing Student Achievement Levels for the National Assessment of Educational Progress (NAEP). Wisconsin appreciates the effort to more clearly define NAEP-Proficient and we welcome more transparency in standard setting procedures. The comments that follow below are specific to draft Principle 3.

Item g, Principle 3
In describing the NAEP-Proficient level, reports shall emphasize that the policy definition is not intended to reflect “grade level” performance expectations, which are typically defined normatively and can vary widely by state and over time. NAEP-Proficient may convey a different meaning from other uses of the term “proficient” in common terminology or in reference to other assessments.

Comment: Wisconsin appreciates the effort to more clearly define for users the meaning of NAEP-Proficient and to differentiate it from other common uses of the term proficient. However, we find that the contrasts with grade level expectations and norms are potentially confusing.

Although the policy achievement level descriptors (ALD) don’t explicitly mention a grade level, the NAEP achievement levels by grade do and are likely to be referenced by users. In addition, in the current assessment and accountability climate, state definitions of proficient are also standards-based and not norm-referenced.

It would be clearer to describe what NAEP-Proficient does mean, rather than, or in addition to, what it does not mean. For instance, it would be helpful to draw the connection between the ALDs, the grade-level NAEP frameworks, and item exemplars.

Item h, Principle 3
An interpretive guide shall accompany NAEP reports, including specific examples of appropriate and inappropriate interpretations and uses of the results.
Comment: We appreciate how comprehensive NAEP reporting is, and that there is a great deal of interpretive information available. However, we agree with the evaluators that the interpretive guidance can sometimes be difficult to find and that some reports may lead users toward inappropriate interpretations of the data.

We recommend that this item be expanded to reflect its importance and to more clearly describe the National Assessment Governing Board’s policy goals related to the means of providing guidance and validating appropriate interpretations.

If you have any questions about our comments, please contact Visalakshi Somasundaram at Visalakshi.Somasundaram@dpi.wi.gov.

Sincerely,

Jennifer Kammerud
Policy Initiatives Advisor

JK:po
I reviewed the ALD policy documents and note the following comments:

- Within the proposed ALDs, the challenge states have with interpretation is in regards to specific content requirements that are assessed in each grade or content area assessment. This challenge is important for state insight as policymakers often wish to leverage NAEP benchmarks in relationship to states’ standard setting process.

- With the pilot exercise noted before the formal standard setting process is undertaken, will any data or impact information be discussed by the pilot panel? It may be helpful to consider potential insight into challenges in interpreting the item map or ordered item book where panelists may struggle with the outcome.

- It may also be valuable for NAEP to consider four proficiency levels and not three. This has been a recent change in many state systems, where the level of not proficient are segmented further to provide insight into weaknesses. Again, if states are often urged to use NAEP reporting levels, this better guides these conversations regarding benchmarks.

Dr. Charity Flores  
_Director, Assessment_  
_Indiana Department of Education_
Dear Members of the NAGB:

As a former member of the National Assessment Governing Board (NAGB), I want to express my support of the proposed changes to the policy regarding "Developing Student Achievement Levels for the National Assessment of Educational Progress." In particular, I appreciate the insertion of the acronym "NAEP" prior to the labels of Basic, Proficient and Advanced. I also fully support the fact that Principle 3, g. describes how "reports shall emphasize that the policy is not intended to reflect "grade level" performance expectations..." While this will not satisfy all critics of the high aspirational nature of the NAEP Proficient level, it will continue to remind the public that it is not equivalent to grade level, especially in light of the fact that "grade level" performance is not defined the same way in all states in this country.

Sincerely,

Louis M. Fabrizio
Chief Executive Officer
Fabrizio Education Consulting, L.L.C.
Alignment of NAEP Achievement Levels

I hope that NAGB and NCES might discontinue reporting student achievement using achievement level percentages.

A recommendation in the mandated 2017 evaluation study suggested that when satisfactory alignment among the frameworks, item pools, achievement-level descriptors, and cut scores in NAEP mathematics and reading has been demonstrated, their designation as trial should be discontinued. However, I do not believe an evaluation of the newly aligned achievement levels will demonstrate them to be reasonable, valid, and informative to the public. They will just be an updated, and fatally flawed example of the percent-above-cut-score metric minus the “trial” designation.

The 2017 evaluation study (page 208) noted that the “percent above cut score metric” is not really useful for following student achievement over time (one of two NAEP stated purposes):

** One of the most common and unwarranted inferences [using achievement level percentages] involves assessing the amount of progress students have made over time, particularly by population groups. For instance, news reports often focus not only on how students are doing at a particular time, but the extent to which the percentage of students scoring Proficient or above has (or has not) improved over successive NAEP assessments. When these comparisons are based on the scale scores, they provide useful information. When they are based on the “percentage Proficient or above” metric and used to compare progress across groups, they can be misleading.

** Holland (2002) focused on misinterpretations associated with using the “percent above a cut score” metric. Although this metric is widely used (for NAEP and many other achievement tests), there are serious limitations to the inferences that [the percent above cut score metric] can support, particularly when evaluating trends over time, gaps among groups, or trends in gaps....

Intended or not, the primary use of NAEP’s achievement level percentages since 2002 has been to accuse state public education officials of racing to the bottom, or dumbing down state tests, or having an “honesty gap” in order to hide the extent that state public schools are failing. I was NAEP State Coordinator for Idaho from 2002 to retirement in 2012. Idaho was always listed as one of the “greatest offenders.” All of this because there were two federal definitions of “proficient,” one by NAEP, and one by NCLB. Moreover, just looking at NAEP: the Proficient Achievement Level was not the same thing as NAEP’s “proficiency in a subject.”

In case, I failed to make my two points. First, achievement level percentages are a poor choice to examine student achievement over time. Moreover, an alignment among the frameworks, item pools, achievement-level descriptors, and cut scores will not change anything. Second, I would like to see NAEP eliminate the main weapon that today’s “school reformers” have used to undermine public confidence in the nation’s public schools.

Bert Stoneberg
this is nonsense spending. nobody in america needs this information. there is no reason to fund this federal dept for this kind of survey. the taxpayers of this country pay for education at these levels:

local
county
state
federal

and we pay for other counties education costs too and military education costs. its time to stop all this spending at the federal level. nobody needs this survey. its just wasteful spending that the fat cat bureaucrats want to make sure they still have a job that never ends. we need to downsize and this can be done away with. it isn't in no way necessary for education. its just fat cat bureaucracy. this comment is for the public record. please receipt. jb ker bk1492@aol.com
To Whom It May Concern:

The California Department of Education (CDE) wishes to express support for the revised policy in Developing Student Achievement Levels for the National Assessment of Education Progress (NAEP). The CDE recognizes many hours of expert, dedicated work in these updated guidelines, and believes the proposed amendments improve and advance the way student achievement on the NAEP is measured and described.

Specifically, the CDE supports:

- The change in terminology from proficient to NAEP proficient – this will help minimize confusion between various assessment standards and clarify the definition of proficiency on NAEP.
- The suggestion that students be referred to as “achieving at a NAEP Basic level,” for example, rather than “NAEP Basic.” This terminology focuses more on the assessment and the shared responsibility of the state, district, school, teacher, student, parents, and community in contributing to the achievement of individual students rather than ascribing qualities to students.
- The Governing Board’s commitment to periodically review the NAEP achievement level descriptions and cut scores. As education standards and curricula change, student achievement will also develop and change. At all levels, educators must continually evaluate what constitutes proficiency, fluency, and mastery of subject matter.

The CDE endorses these changes and appreciates the opportunity to comment. Please contact Julie Williams at 916-319-0408 or by email at julwilli@cde.ca.gov if you wish to discuss these comments further.

Sincerely,

/s/

Michelle Center, Director
Assessment Development and Administration Division

MC:jw
Comments on draft NAGB revised policy on Developing Student Achievement Levels for the National Assessment of Educational Progress (NAEP).


September 22, 2018

I appreciate the opportunity to comment on the draft policy. It was on my watch as (first) NAGB chairman that achievement levels were first developed, and we now have more than a quarter century of experience with them.

Though they’ve been endlessly criticized (mostly for being too high, or not being scientific enough, or not having proven validity), they’ve also become indispensable to American K-12 education, widely used, cited, calibrated with, and depended upon in so many ways.

NAEP’s achievement levels are, in effect, the only rigorous and revealing national standards or benchmarks that we have for K-12 education, at least the only ones that are universal (unlike SAT/ACT), that yield state-level (and in the case of TUDA municipal-level) data (unlike PISA, for example), and that apply to the elementary and middle grades as well as high school (unlike AP, for example).

Vitally, they yield comparative data—from state to state, state to nation, sometimes city to city and city to state and nation—and they also yield trend data so that changes over time in student achievement can be tracked. Sure, there are other ways of doing that (e.g. ups and downs in scale scores for the top and bottom deciles) but they’re basically beyond the easy understanding of parents, most policymakers and many educators. Achievement levels are crystal clear, readily intelligible.

The ability to compare NAEP results across jurisdictions with the help of achievement levels is more important than ever now that the PARCC and Smarter Balanced assessments are more or less imploding, with almost every state—even those with common standards—devising and deploying its own “branded” assessment. Had those coalitions stuck together, we would have had other sources of inter-state comparison but, sadly, we don’t.

Are achievement levels set too high, as some allege (e.g. the “How High the Bar” attack on both Common Core standards and NAEP achievement levels)? What so many critics refuse to acknowledge is that the achievement levels—at least the proficient and advanced levels—are aspirational, and meant to be, much like scoring 4 or 5 on PARCC, or 4 or 5 on an AP exam. (PARCC is on my mind because we’re still using it in Maryland where I’m on the state board of education.) Everybody knows that most kids aren’t there yet, but there is where they ought to be, and where the education system should
be doing its utmost to get them. (It's also where many, many more kids are in other
countries, including many kids who are as “diverse” as young Americans. We know this
in part from successful efforts to equate NAEP and PISA.)

The “too high a bar” criticism was pretty much knocked out of the water by NAGB’s own
careful research demonstrating (to oversimplify a bit) that 12th grade NAEP proficient is
approximately equal to true college readiness. That’s about as solid a validity check as
we’re ever going to see, at least for the “proficient” level! (I strongly suspect that NAEP
“advanced” could be linked to PISA 5/6—research worth undertaking.)

The other issue that keeps arising—and it’s more visible than I think desirable in the
new draft policy statement—is the ceaseless desire of technical folks to turn
achievement level-setting into a technical or scientific process. At its heart, it isn’t and
cannot be, for at its heart it’s a judgmental statement of aspiration. Aspirations are
almost never “scientific” and while judgments can and should be informed by technical
analysis, in the end they’re judgments, not “findings.”

My only real criticism of the new draft is that it’s not nearly as clear as the old one on
this matter of judgment. It alludes to it, but in a semi-apologetic way, instead of bluntly
and candidly declaring that it’s NAGB’s job to exercise judgment and nowhere is that
more important than in determining the “cut points” for the achievement levels. The
reason there’s a Governing Board, comprised as it is, is to make such judgments! You
can conjure up all the complex procedures, analyses and technical advisors you like, but
in the end they cannot transform an act of judgment into science, and you shouldn’t
even hint that they might!

Having said that, technical expertise is valuable in a number of ways such as, for
example, keeping achievement levels calibrated from assessment to assessment, else
the trend information would be useless. You don’t want your yardstick bending,
stretching or shrinking! In the end, however, just as a judge and jury hear from all sorts
of expert witnesses, it’s judgment that they must exercise. It’s judgment that NAGB
must exercise, and I wish the revised statement of principles were clearer on that point
and less self-conscious about all the trappings of process, technique, and expertise.

One last point. In my experience (both on NAGB, as assistant secretary of education
and in sundry other roles), the non-educator members of panels such as you convene
for standard-setting are arguably more valuable than the educators, for they have a
better sense of the level of performance that the “real world” wants to see from kids
emerging from the K-12 system. Employers, newspaper editors, instructors of first-year
college students, authors of articles and blogs, military recruiters—I could go on. Please
don’t stint on incorporating a good cross-section of them on the level-setting panels!

Thanks again for this opportunity to weigh in.
Dear Ms. Rosenberg and members of NAGB:
Thank you for very much for the opportunity you are providing to the draft policy regarding achievement levels.
As a former board member for NAGB representing the role of a public high school principal from Colorado (at that time), I feel compelled to respond to the ongoing debate about whether achievement levels are benchmarks that are too high or not useful or valid.
I was fortunate to serve on the Achievement Level Committee and remember well the rigorous discussions and debates with math experts, researchers and a wide range of faculty from across the country. The need for well-prepared students of all backgrounds and income levels was well documented then, as it is now. All students in public, private or faith-based schools all across the country deserve to be challenged and supported to reach their highest potential. Having worked in public education during my k-12 and now higher education career, I know how much schools care and challenge students.
Achievement Levels are and have been about the educational aspirations we have for all students, about our competitiveness, about our talent development to build the workforce and economic and cultural development worthy of the human spirit.
I urge NAGB to stay the course, to continue to challenge our students, our schools and indeed our country to the highest levels of achievement. And I urge NAGB to continue to refine and improve assessments as new technologies and protocols are developed—however to NOT give up on the need of high and rigorous standards and achievement levels.
I can tell you that in the last 15 years in the community college sector, I see on daily basis the need for our students to increase their mathematical skill—and their persistence is paying off as they prepare for careers and professions in STEM and the challenges of our technological society.
Please feel free to call on me—I would be happy to say more.
Sincerely,

Christine Johnson, PhD
Chancellor
Community Colleges of Spokane
501 N. Riverpoint Blvd. Suite 110
PO Box 6000, MS 1001
Spokane, WA 99217
Concerning the Achievement Level Setting Program,

I am writing today to express my urgent plea that the current NAEP Achievement Level system be restructured and renamed effective immediately. For years, the Levels have confused, confounded, and misled the American Public, the media, politicians, and policy makers to the detriment of our nations schools. Teachers have been fired and schools have closed because of the misuse of NAEP data and its leveling system. America’s children have been made out to be ‘failing’ when they score below Proficient. when in reality the passing mark is out of reach and always will be. The National Center for Education Statistics (NCES) has clearly stated that ‘proficient’ is not synonymous with grade level performance. But when a metric is so clearly misused, misunderstood, and abused it is clearly time for an immediate restructuring. That time is now.

Emily Maurek
Teacher
September 29, 2018

National Assessment Governing Board
800 North Capitol Street, N.W., Suite 825
Washington, DC 20002

Dear Governing Board Members,

Thank you for the opportunity to provide feedback regarding your draft policy, “Developing Student Achievement Levels for the National Assessment of Educational Progress.”

On behalf of the trustees of the Thomas B. Fordham Institute, I would like to commend the draft for its clarity and rigor. NAEP’s achievement levels have served a vital purpose for almost thirty years, clearly communicating to the nation the proportion of students who achieve at a basic, proficient, or advanced level, and setting a high, aspirational standard for the nation’s schools and educators.

The revisions to the previous policy that are under consideration—such as clarifications regarding the meaning of the various performance levels and guidelines for how they are communicated to the public and the press—are measured, and appropriately so.

It would be an enormous mistake to follow the advice of some advocates who want NAGB to describe the proficient level as “extremely demanding” and to equate the basic level with “roughly analogous to performance at grade level.” This would represent a lowering of the bar, plain and simple.

To be sure, the proficient standard is a high one—indicating “solid” performance and competency in “challenging” material, in the words of your draft policy. It would indeed be foolish to expect one hundred percent of students—in the United States or anywhere—to be able to meet such a high bar. And it’s unfair when some critics of the nation’s schools imply that anything less than universal proficiency is tantamount to failure.

Yet, as NAGB’s own studies have demonstrated, the proficient standard is not much, if any, higher than the aspirations of America’s moms and dads. Research from Learning Heroes and others indicates that almost all parents expect their children to go to college, and most children share those aspirations as well. Setting aside whether we actually need everyone to go to college (or even attend “postsecondary education”), it’s undeniable that higher education is now part of the American Dream. And NAGB has determined that in reading at least, twelfth graders who are proficient are also well-prepared for college. (For math, the preparedness level is between basic and proficient.)

It makes sense, then, for proficient to be the goal, and for Americans to continue to receive reports regarding how many students are attaining it.

Happily, most state policymakers appear to agree. Analyses by Education Next and other indicate that in recent years states have dramatically raised the rigor of their own annual assessments, so that
“proficient” on state tests is now much closer to “proficient” on NAEP. In some cases it’s even higher. This is evidence, from America’s laboratories of democracy, that “proficient” is a reasonable if challenging goal.

And again, it would be inappropriate for states to consider schools that do not get 100 percent of students to the proficient level to be failures. But there is comforting news on that front as well; under the Every Student Succeeds Act, the majority of states have moved aggressively to weight student progress over time much more heavily than the proportion of students attaining proficiency. That too indicates that proficiency is seen as an aspirational goal—one that may never be attained by everyone, but one that is a marker of solid performance.

Your draft policy reflects a good-faith effort to be responsive to the valid concerns made about NAEP’s standard-setting processes, without lowering the bar. Please don’t make significant changes that would diminish this reasonable compromise.

Cordially,

Michael J. Petrilli
President
Thomas B. Fordham Institute

CC: David Driscoll, chair, Thomas B. Fordham Institute Board of Trustees
    Chester E. Finn, Jr., president emeritus, Thomas B. Fordham Institute
I have just one point - but a crucial one.  
Performance standards that are describe linguistically are inevitably highly ambiguous.  
Rich exemplification of:  
* the types of task students will be asked to perform  
* how performances will be scored  
* the presumed degree of unfamiliarity of the task (not the genre, of course) makes the standards clear.  

I and my Shell Centre colleagues would be happy to exemplify such exemplification!  
It also has the effect of providing learning goals for the classroom.  

Hugh Burkhardt  
Mathematics Assessment Project  
map.mathshell.org  
see also  
https://www.mathunion.org/icmi/awards/past-recipients/2016-icmi-award-winners
The National Assessment Governing Board should change the achievement benchmarks for the National Assessment of Educational Progress. These standards and benchmarks must be expressed in terms that parents and the public can understand.

The current achievement levels do not do that. Those levels, instead, have led to confusion in the news media and among teachers, parents, and the general public. It appears that very high aspirations ruled their development, instead of realistic conclusions based on sound data.

The consequences of the current approach are shown in recent research linking NAEP’s benchmarks to international assessments: the majority of students in most nations cannot clear NAEP’s proficiency bar. A disregard over many years of such studies done by respectable organizations showing the deficiencies in the processes used to develop these achievement benchmarks is the reason for these unrealistic results.

The National Assessment Governing Board has been a good steward of NAEP, with this major exception. I urge NAGB to establish a sound process to rewrite these achievement benchmarks. In addition, the public should not only be given more time to comment on these measures, but the advice given should also be considered with an open mind. I would expect nothing less from the Board.

I realize that this message is being sent after the [September 30] deadline due to my being out of the country for a while. I feel so strongly about this issue that I am sending this anyway. Thank you.

Jack Jennings
this project should not be funded. we need a much smaller federal dept of educaton. this dept is too large, it is a spendthrift with no validity for the excessive trillions of dollars it spends. it is letting students take out huge loans for nothing at all. they are never productive. they never intend to pay back. they never will pay back. the entire system is out of control. it sucks. this agency deserves an f minus for its work. it needs to be shut down. we can do without 4 levels of education depts.

we pay for local, county, state and federal levels for education. no other country in the world requires 4 different govt levels for education. they do a better job with 1 or 2 levels. f minus for this dept and its corruption, its fast cat bureaucracy. nothing good comes from the pedantry bureaucracy of this dept. this comment is for the public record. please receipt. jean public jean public1@yahoo.com
I strongly encourage you to set "Proficient" for grade 12 at the level required to earn a "B" in entry-level credit-bearing courses at moderately competitive American colleges and universities. This standard could be set by examining student work, assigned texts, and assessments at a sampling of these universities/courses. This would make standard setting more objective and less political and would provide clarity so that the average American high school student and parent could understand the results.

Kimberly Fleming, PhD  
Subscribe to Core Education's Newsletter  
www.CoreEducationLLC.com  
Blog: www.CoreEducationLLC.com/blog  
Twitter: @CoreEducation1
The following pieces constitute Marc Tucker, President and CEO of the National Center on Education and the Economy’s comments for the National Assessment Governing Board’s policy statement on Student Achievement Levels for the National Assessment of Educational Progress (NAEP). They originally appeared in his Education Week blog, Top Performers.

Blog #365

NAEP Must Get Its Standards Right This Time

Forgive me if I am a little cynical about the eternal dance between measurement and accountability when it comes to reporting on the progress and achievement of American school children. From the beginning, the leaders of our state education systems have invited testing experts to help them set the cut points for passing or not passing the state tests. They listen gravely to the advice of the experts, then ask them how many students will fail at the recommended cut point and set a new one at a point that is politically tolerable.

The heads of municipal school systems for a long time picked the test they would use to report student performance from vendors who offered to compare the performance of their students to that of any of many different student bodies elsewhere. The superintendent would pick the comparison that would make their district look the best. All the insiders knew that was how it worked. Only the public was fooled.

When George W. Bush became President, he wanted to hold every state, district and school to a common national standard. He couldn’t get what he wanted, but he did the next best thing. He required all the states to participate in the National Assessment of Educational Progress, “The Nation’s Report Card.” For the first time, the performance of all the states could be compared on a common metric.

Well, that was interesting. President Bush’s signature education program, No Child Left Behind, required each state to set its own standards for student performance and then commit to reaching those standards by 2014. These standards, of course, could be different. It turned out, rather famously, that the states claiming to make the most progress toward reaching their standards were those that performed the worst on NAEP. The governors of the states that ended up with egg on their face were among the strongest supporters of the development of the Common Core State Standards. They did not want to be embarrassed that way again.

But, as we all know now, not all state versions of the Common Core are the same and there are a number of states that have not embraced the Common Core in any form. And the state consortia formed to create common assessments of the Common Core have withered on the vine, so there is now no prospect that the Common Core and its associated tests will enable all schools, districts and states to compare their performance to all the others on a common, honest metric.
That leaves NAEP. The way NAEP is done permits the observer to compare scores among states for the grade levels and subjects it assesses. But what do those scores mean? In an attempt to answer that question, the NAEP Governing Board, starting a quarter century ago, settled on three distinct levels of performance: NAEP Basic, NAEP Proficient and NAEP Advanced. The Board’s policy statements define those terms and describe the process by which the Board will decide on the cut scores that demarcate the boundaries between performance levels.

What makes this moment special is that NAEP is now, today, engaged in the first major revision of the procedures by which those standards are set. You have a chance to weigh in. Look here to see how you can make your views known to the Board.

This policy is really important, because the views that Americans have about the performance of their schools are significantly affected by the generous press attention that the NAEP reports routinely get. But what those reports mean has a lot to do with how NAEP defines performance. That is what this blog is about.

Let’s take the meaning of the word “proficient.” The new draft standards say proficient means “…solid academic performance for each NAEP assessment. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real world situations, and analytical skills appropriate to the subject matter.”

Hmmmm… What does “solid” mean? Who defines what it means to be “competent”? What is terribly “challenging” to one child might be super easy for another. This definition is quite obviously a matter of judgment. And that is the way the issue is treated by the draft. It says a panel of “subject matter experts will be convened to recommend achievement level cut scores....” What really counts here is their opinion.

And then, of course, it says that these subject matter experts do not decide on the cut scores, but instead make recommendations to the full NAEP Board. It explicitly directs that the Board have information on the effects of setting the cut scores at different levels—that is, how many students are likely to be found proficient. That strongly suggests that political judgment will play a decisive role in cut score setting, just as it has always done at the state level.

But then the document says that these judgments need to be “valid.” You would think that would mean coming up with empirical data showing that students said to be proficient actually are, in some commonsense meaning of the word, proficient. But it does not mean that. It means that there is empirical evidence that students who are said to be proficient do in fact have the capacities specified in the definition. But that is circular. What do you mean by challenging? Answer: “Whatever I have measured.”

What would make it uncircular? Answer: Knowing whether a student is proficient or not would have some meaning for me if I knew whether that student could do something in particular that is important to me or the student. For example, whether that student is ready for college or
ready for a career. Now, you say, fine, I can go with that, but can you?

Ready for which college? The fly-by-night “institution” down the street that offers no instruction but is ready to take your college loan money this afternoon or Michigan State? Ready for which career? A career as a cashier in a fast food chain or a career in finance on Wall Street? The Department of Education’s draft provides no guidance on any of these points.

Nor does the history here give this reader confidence that the process described will get this right unless such guidance is provided. The NAEP Governing Board has done a good job of sponsoring research that correlates scores on NAEP with certain college outcomes and workers’ incomes. But that does nothing to tell students, parents or teachers or even state policy makers what students need to study or how well they have to do to be college and career ready. If the members of the Board think that one can only be proficient in mathematics at the 12th-grade level if one has demonstrated a thorough command of the topics typically included in an Algebra II course, then the people who construct the test will include a lot of Algebra II questions on the test and policy makers will tell the schools everyone has to take Algebra II. If the Governing Board says that, in their judgement, that is what proficient means, who is to say they are wrong?

Actually, me. One does not have to have mastered the content of the typical Algebra II course to succeed in college or career. This is how I know that:

First, a very large fraction of high school students going to college in the United State either do all their college work in a community college or take their first two years of a four-year college program in a community college.

Second, the nation’s primary provider of career training, meaning vocational education and training, is the nation’s community colleges.

Third, successfully completing the first year of a typical community college program is a good predictor of the likelihood that the individual will successfully complete a two-year degree program or acquire an occupational certificate of value to an employer.

Fourth, it follows from “1,” “2” and “3” above that, if one cannot succeed in the first year of a typical community college program, one’s chances of succeeding in further college or career are slim, and the converse is also true. So, one could reasonably say that whether or not the high school student is ready for success in the first year of a typical community college program is a very good measure of the degree to which the student is “ready for college and career.” It does not mean the students has a high probability of success in the first-year program at Stanford or in the first year of a program designed to train medical technicians to administer and interpret sonograms, but it does specify a standard of proficiency that is specific and broadly applicable, a standard that would have intuitive appeal to millions of American students, parents and college admissions officers. In most advanced industrial countries, there are one or more high school leaving credentials that are matched to the requirements for going
on to university or to advanced occupational training. Would it not make sense for “proficient” to mean just that, using the demands of credit-bearing courses in the first year of our community colleges as the benchmark standard? To do this right, the National Assessment Governing Board would have to know not just what cut score to use on a general test of mathematics, but what topics in mathematics would have to be mastered to what level to enable the student to succeed in College Mathematics or College Algebra.

Our organization has done the research needed to establish the benchmarks for such a standard, at least with respect to reading, writing and numeracy. We did not do it the way NAEP has done it, by asking “experts” what they think the standard ought to be. Industrial psychologists found out years ago that that approach almost never actually ends up with descriptions of what is required to do a particular job or the level and kind of education and training needed to do that job. Instead they study the job itself, the way real people do it and then use that information to figure out what sort of education and training they need.

That’s what we did. We gathered the most widely used texts used in the most commonly taken initial credit-bearing courses in a randomly selected set of community colleges and asked leading reading experts to determine their reading level. We asked for graded writing samples from typical assignments given to students along with scored exams and had those reviewed by leading writing experts. And we reviewed the texts for the courses called College Mathematics and College Algebra and had them reviewed by the nation’s leading mathematics experts.

It turns out that College Mathematics and College Algebra are mostly topics covered in Algebra I, and a little geometry, statistics and probability. Students leaving high school do not need to be proficient in Algebra II in order to study Algebra I. The first-year texts are mostly written at the 12th-grade level, but a large fraction of our high school graduates cannot comprehend what is written in them. Many of the community college instructors told us that they do not assign writing to their students because the students cannot write and the instructors do not think they were hired to teach basic writing. Yes, some of the NAEP benchmark standards for mathematics are well above any reasonable definition of “proficient.” But a standard of proficiency that was based on what it would actually take to succeed in the first year of community college would be way below the global standard for college level work in the advanced industrialized nations.

If I were on the NAEP Board, I would press for setting proficiency standards based on what empirical data—not anyone’s “expert opinion”—tell us about the content and performance requirements for success in the first year of the typical community college. I would urge my fellow board members to adopt a policy for reporting to the American people on how many high school students reach that standard at the end of high school and how many students are on a trajectory to reach that standard in elementary and middle school. I would push NAEP to tell the American people that this benchmark should be used by the states to set a target for what their students should achieve by the end of tenth grade, because that would represent a level of achievement for students of that age comparable to the level achieved by most
students in the top-performing countries by that time and there is no reason why we should expect less.

And lastly, and most importantly, I would tell them that NAEP is the last redoubt, the last remaining hope that the United States will have an instrument that we can use to get an honest measure of how our students are doing. If we lose that check, if anyone can say whatever they like about how our students are doing, ignorance will not be bliss.

Blog #366

Setting NAEP Performance Standards: How We Got Here

Last week, I wrote a blog commenting on the draft policy on performance standards recently issued by the National Assessment Governing Board. In it, I called for performance standards based not on people’s opinions about what constitutes basic, proficient and advanced performance at three grade levels but on what the evidence shows it takes to succeed in the first-year program of a typical community college, on the grounds that, for at least half of our high school graduates, our community colleges are the gateway to both careers requiring occupational certificates and to two-year and four-year college degrees.

Among the responses I got was one from Jim Pellegrino, who sent me two papers he thought might interest me. One was the chapter on setting performance standards in a book that he, Lee Jones and Karen Mitchell had done for the National Academies in 2009 titled Grading the Nation’s Report Card: Evaluating NAEP and Transforming the Assessment of Educational Progress. The other was a paper by Albert Beaton, Robert Linn and George Bohrnstedt written for the National Center for Educational Statistics in 2012 titled Alternative Approaches to Setting Performance Standards for the National Assessment of Educational Progress (NAEP). They make for very interesting reading for those of us who care about the usefulness of The Nation’s Report Card.

Reports from the National Academies are usually written in the most measured scholarly language imaginable. Not this one. The smoke rises from its pages. NAEP’s standard-setting process, it says, is “fundamentally flawed.”

Reviewing critiques of the NAEP standard-setting process offered through the 1990s, the chapter from the National Academies book quotes analysts who described the process of making judgements a “nearly impossible task” for the raters, pointed out that the process produced different cut scores for different kinds of test items (e.g., open ended vs. multiple choice) and said the cut scores had been set at levels that were simply not credible when compared to evidence from other well-regarded assessments.

All these issues came to a head with the 1996 science assessment. The results showed that, at all three grade levels, very low percentages of students scored proficient, and, at the high
school level, hardly any students had made it into the advanced level. The obvious conclusion was that students who had earned good grades on the Advanced Placement tests in science were not considered by NAEP to have achieved at advanced levels. One would also have to conclude that students who had done very well on the Trends in International Mathematics and Science Study (TIMSS) were not doing advanced work either.

These results were not all that unusual. In other cases, too, the NAEP findings on the performance levels of American students seemed to be way off in both directions. The standards were just not credible.

So the Board adjusted the cut scores to make them more credible. But then the critics noticed that the new cut scores did not match up with the performance descriptions for the standards. What was defined as proficient work in the descriptions was not what was tested by the items used in the proficient range. So the Board took out the definitions! Well that was one way to deal with the discrepancy. Then one could say a student was proficient if he or she scored in a certain score range, but one could no longer say what that meant in terms of what the student could do.

The Board empaneled another group to write new descriptions of what the performance levels meant. This time they wrote descriptions not based on what a student should know to be proficient, but rather on descriptions of what they currently know and can do. “...[I]nstead of reporting achievement results relative to an established standard of performance... the science report presented results that were based on NAGB’s a priori judgment as to what constituted reasonable percentages of students at the three achievement levels.”

The problem with that, of course, is that the public might reasonably think that a student who was rated proficient on a subject at a certain grade level by NAEP was able to do what a student needed to do to be successful according to expert judgment, but it did not mean any such thing.

What had really happened was that a complex technical process that was supposed to produce findings about student achievement against common sense standards had failed badly. The process had produced findings showing that, in some cases, the standards were much too high by any reasonable measure and in other cases much too low. In the end, the NAEP Board did what such bodies had always done before in such cases. It adjusted the results to produce a politically palatable result without a solid rationale for its decision. It had been and was still the case that it was very unclear what the performance standards were or what they ought to be.

The problem, as I pointed out in my last blog, is that the nation was much more focused on finding accurate, unfudgeable measures of student performance with which to measure the performance of state education systems, districts and schools than ever before. The issue of what the performance standards meant and how they should be developed would not go away.
You might reasonably assume that a slashing attack like this from such a distinguished group of critics would have led the NAEP Governing Board to respond to the critique with a fix for its standard-setting problems. But that did not happen. Years later, the process for setting the standards had not changed substantially, so another, no-less-distinguished band of scholars took up where Pellegrino and his colleagues had left off. Beaton, Linn and Bohrnstedt were members of the NAEP Validity Studies Panel, now chaired by George Bohrnstedt himself. Their paper was written as part of that program of studies.

In it the authors explore three possible alternatives to the process for setting NAEP achievement levels I described above.

The first alternative would be to make the achievement cut points predictive. The cut point for the end of elementary school assessment would predict the likelihood of success in middle school, the cut point for the end of middle school assessment would predict the likelihood of student success in high school, and the purpose of the end of high school assessment would be to predict the likelihood of success in college and career.

The second alternative would be to “benchmark the achievement levels against international standards.”

Their third alternative was to use percentile rankings to set base-year norms against which progress could be measured in succeeding years. The authors acknowledge that hybrids of these approaches could be developed, too.

I was astonished when I read this paper. Taken together, its proposals mirror the plan I described, in more detail, in last week’s blog. But those ideas are not new. I first proposed them years ago when my organization created a program called Excellence for All to put them to the test in the field in several states including Kentucky, Arizona and Mississippi. The assessments we used were not the NAEP assessments—they were not designed to be used as census assessments of all the students in a school—but the International General Certificate of Secondary Education exams offered by Cambridge Assessment International Education of Cambridge, England.

The idea was to set a high standard for all high school students based on what it would take to succeed in the first year of a typical community college program and design the program of the high school so that most students would reach that standard by the end of grade 10 and almost all would get there by the time they graduated high school. Students who reached the standard by the end of grade 10 would be able to enroll in a demanding upper division high school program like International Baccalaureate, a whole program of AP courses or the Cambridge Diploma program, all of which are designed to qualify students who get good grades on those programs into the world’s most selective colleges and universities. Or they could take a full program of community college courses and wind up either with a strong vocational credential or two years of college credit, ready to transfer at the end of high school straight into the junior year of a state college or university.
We needed a team of top education researchers to help us set the right pass points on the Cambridge exams to make this program work. We asked Jim Pellegrino and Howard Everson to chair the Technical Advisory Committee. The other members were Catherine Snow, Phil Daro, Bob Linn, Richard Duran, Ed Haertel, Dylan Wiliam, Joan Herman and Lloyd Bond.

Beaton, Linn and Bohrnstedt recommend that NAEP do an empirical study to benchmark college and career readiness. We did that, with a research plan approved by the members of this Technical Advisory Committee (TAC) and drawing on the services of several of its members. Beaton, Linn and Bohrnstedt recommended that the NAEP high school performance standards be set to predict the likelihood of student success against the empirically determined college and career benchmark. We did that, using the Cambridge exams, rather than the NAEP assessments, also under the supervision of the members of our TAC and using their services. Beaton, Linn and Bohrnstedt recommended that NAEP benchmark international student performance standards and design NAEP performance standards to predict student success on those benchmarks, too. We did that, too, using instead of the NAEP assessments the Cambridge assessments.

The only difference between what we did and what was recommended in the paper was that we substituted Cambridge examinations for NAEP assessments. There is no reason why NAEP could not replicate what we did substituting NAEP assessments for Cambridge assessments.

Not only that, but, because we have already done the work, we know what an empirical study of college and career readiness, defined as readiness for success in the first year of a typical community college program, requires. We also know that this is pretty much the same performance standard that is met by the typical student who is entering gymnasium in Europe or beginning to take “A”-level exams in England.

The point is that a very good model exists and has been tested for producing performance standards for NAEP that would address and resolve the problems in setting performance levels that have dogged NAEP for years.

This is no esoteric matter. For all the reasons I advanced in last week’s blog, NAEP is this country’s last redoubt for honest comparisons of the performance of a state’s education system to that of other states. It is certainly true that performance could just be reported as a number on a scale. But what does that number mean? What does it say about what the students know and can do? About whether they are ready for the next stage of their education or to begin a rewarding career? How their performance compares to the performance of students of the same age in other countries? If you care about the answers to these questions, you should care about the way NAEP sets its performance standards.

The changes now proposed by NAEP to the standard setting process do not address these issues. If you think the issues I have raised are important, write to the NAEP Board with your comments.
Comments on NAEP Standard Setting

The draft policy statement on NAEP standard setting is clear and generally reflective of current best practices. I offer the following comments and suggestions for the Board’s consideration.

p. 6: There shall be no content ALDs developed for performance below the NAEP Basic level.

If the only purpose of the ALDs is to guide standard setting panelists in setting cut scores, it would be reasonable not to have ALDs below the Basic level. However, since ALDs may also be used to guide item development, via the creation of range ALDs, there should be at least some indication of what Below Basic looks like so that items assessing the knowledge and skills of students performing at that level could be created and matched to an appropriate range ALD.

p. 7: h) The process shall consist of at least two achievement level setting meetings with distinct groups of panelists, a pilot study, and an operational meeting.

This is a great idea. Experience has shown that in high-stakes situations such as NAEP, a pilot study eliminates all manner of bugs that could appear in the operational standard setting. The Board has the resources to conduct split-group meetings and should do so. Individual standard setting plans should indicate how cut scores will be compared/resolved/aggregated across groups.

p. 8: c)...The Board shall revisit and may revise content ALDs following the achievement level setting to ensure that they are consistent with empirical evidence of student performance.

Modifying ALDs after standard setting is a time-honored practice but a dangerous one. If the content ALDs are sufficiently vetted, and if the range ALDs have been properly developed, vetted, and applied in item development, and if threshold ALDs faithfully reflect the intent of the range ALDs, then any adjustment at the end of standard setting should be to align cut scores with fixed ALDs, not the other way around. By the time standard setting is conducted, ALDs will be highly visible and will have taken on a life of their own. Modifying them will be much more visible than modifying cut scores. If necessary, reporting ALDs may be fine tuned to reflect specific outcomes of standard setting, so long as they can be legitimately aligned to the policy and threshold ALDs. Please consider revising this portion of the plan.

p. 8: d) The Board shall examine and consider all evidence related to validity of the achievement level setting activities.

In many situations, cut score validation now includes review of ALDs and cut scores by content matter experts from higher grades (e.g., middle school educators for grade 4 standards, postsecondary educators for grade 12 standards) as well as predictive validity with respect to performance in subsequent grades. The Board may wish to consider extending validation of NAEP cut scores by including on the front end a review of all ALDs, test items, test forms, and
cut scores by educators beyond the grades for which the NAEP forms are designed and on the back end predictive validity studies using school performance in subsequent grades (e.g., relevant course grades for fifth graders who took the fourth grade exam the previous year). [See also Principle 4, section d.]

General

Principle 4 sets forth excellent guidance for ongoing validation and updating of cut scores and integration of standard setting with all other aspects of test development, administration, and reporting.

Thank you for the opportunity to participate in this review.

Michael B. Bunch, Ph.D.
Senior Vice President
Measurement Incorporated
I helped write Oregon's academic standards and standardized tests in the 90's, based on national and international expectations.

In addition, I was a guest teacher in Europe where I learned about their national standards, the PISA and exams for college prep seniors.

I can tell you right now that the current US standards based on the NAEP are unwieldy, inconsistent and completely unrealistic for students who aren't college bound.

No other nation requires all students to complete Algebra II. That alone is a major reason for America's high dropout rate. I know teachers forced to teach special education students to mindlessly copy quadratic equations just to prove the entire school is working toward the Common Core and NAEP standards.

Finally, setting the standards at NAEP's mastery level instead of at grade level creates a generation of kids that feel like failures. Regrettably, I was part of the creation of the rubric scale of 1-4 designed to match the GPA system, whereby 3's or B's equate with on-track for college. But internationally the mastery level is only expected of the top third of students, not everyone.

Let's acknowledge that the NAEP holds up an ideal unattainable by the average student under the best of circumstances. Let's return US standards to grade level expectations, whereby average students will be rewarded for passing the tests based on those standards and the stars will always shine with A's and B's.

Rachel Rich
Retired High School Teacher
Past President of OATG
Past Board Member of PNCFL, COFLT, AATG
Goethe Institute Teacher Trainer
Workshop Presenter
Dear Sirs:

I offer the following information as you consider changes to the NAEP Achievement Level scoring system.

Based on my comparison of NAEP data to data from ACT Grade 8 EXPLORE testing conducted in Kentucky from 2007 to 2015 (See Attachment 1), it appears the current NAEP “Proficient” achievement level is already well aligned to identify students on track as of the eighth grade to be ready for college upon high school graduation. Therefore, I encourage NAGB to make no changes to the scoring for Proficient on NAEP.

Sincerely yours,

(Signed)

Richard G. Innes

1 Atch
Comparison of Kentucky's Benchmark Score Results from ACT's EXPLORE with NAEP Grade 8 Proficiency Rates
Attachment 1
Comparison of Kentucky's Benchmark Score Results from ACT's EXPLORE with NAEP Grade 8 Proficiency Rates

This research takes advantage of same cohort testing in Kentucky with the NAEP and the ACT Incorporated’s EXPLORE assessments to explore the relationship between the percentage of students scoring at or above the NAEP "Proficient" level and the percentage of students scoring at or above the readiness "Benchmark" scores on the EXPLORE assessments.

The Kentucky wide data is available for years from 2007 to 2015.

Because the Jefferson County Public School District in Kentucky also participated in the NAEP Trial Urban District Assessments from 2009 onward, an additional set of NAEP to EXPLORE comparison data is available.

For those who are not familiar with the EXPLORE, the ACT, Incorporated expended considerable effort to insure the EXPLORE test Benchmark Scores are closely linked to the College Readiness Benchmark Scores for the ACT college entrance test. Those ACT test Benchmark Scores are the result of empirical studies that relate those scores to the real performance of college freshmen in related college courses. Thus, the EXPLORE has a strong relationship to what students need for college and careers as of the eighth grade.

Data Sources:

The Kentucky and Jefferson County Public Schools EXPLORE Benchmark results are available in an Excel Spreadsheet available here.


NAEP percentages of students scoring at or above "Proficient" in math and reading for years prior to 2015 were obtained from the NAEP Data Explorer on June 5, 2015. The NAEP 2015 data were obtained from the NAEP Report Cards supporting Excel spreadsheets on October 28, 2015.

Comparisons of NAEP and EXPLORE Results

The four figures below compare the percentage of students scoring at or above NAEP Proficient to the percentage of the same cohort of students scoring at or above the EXPLORE Readiness Benchmark Score. Note the close agreement throughout.
Figure 1 – Kentucky Reading

Kentucky READING
Percentage of Kentucky 8th Grade Students Scoring Proficient or Above on NAEP Vs. Percentage of Same Student Cohort Scoring At or Above The EXPLORE Benchmark Score, 2006-07 to 2014-15 School Terms

Figure 2 – Kentucky Math

Kentucky Math
Percentage of Kentucky 8th Grade Students Scoring Proficient or Above on NAEP Vs. Percentage of Same Student Cohort Scoring At or Above The EXPLORE Benchmark Score, 2006-07 to 2014-15 School Terms

Innes to NAGB on NAEP Achievement Level Scoring
Figure 3 – Jefferson County Reading

Jefferson County READING
Percentage of Jefferson County Public Schools 8th Grade Students Scoring Proficient or Above on NAEP Vs. Percentage of Same Student Cohort Scoring At or Above The EXPLORE Benchmark Score 2008-09 to 2014-15 School Terms

Figure 4 – Jefferson County Math

Jefferson County MATH
Percentage of Jefferson County Public Schools 8th Grade Students Scoring Proficient or Above on NAEP Vs. Percentage of Same Student Cohort Scoring At or Above The EXPLORE Benchmark Score 2008-09 to 2014-15 School Terms
Observations on NAEP to EXPLORE - Math

The Kentucky NAEP to EXPLORE math agreement shows only a 1-point differential in 2007 and 2011, a 2-point spread in 2009 and a 4-point spread in 2013 and 2015.

In Jefferson County, the agreement was exact in both 2009 and 2011 and differed by only 1 point in 2013. The Jefferson County difference grew to three points in math in 2015.

Observations on NAEP to EXPLORE - Reading

The Kentucky NAEP to EXPLORE reading difference is 7 points in 2007, but this was reduced to only a 2-point spread in the 2009 and increased slightly to a 3-point difference in 2011 and a 4-point difference in 2013 and 2015.

In Jefferson County, the agreement in 2009 was also quite close with just a 3-point differential, which increased to a 4-point difference in 2011 and then reduced to a 3-point spread in 2013, finally becoming a perfect tie in 2015.

Implications

The close agreement for NAEP and EXPLORE for both math and reading is remarkable given the fact that the development process for these two assessments and their scoring schemes are quite different.

Also, there is sampling error in the NAEP. After that sampling error is considered, most scores reported here from the EXPLORE are essentially equivalent to the NAEP.

The findings here help further inform long-standing concerns about the accuracy and meaning of the NAEP Achievement Level Scores, at least for the meaning of NAEP "Proficient" determinations at the eighth grade level in math and reading.

If a good psychometric case can be made that the NAEP 4th grade achievement level scores are well-linked to the eighth grade math and reading scores, then we would also have the ability to reach down into elementary grades to gain insight into how many students in each state are on track at that early level to be college and career ready in those subjects. That is obviously a valuable piece of information.

It would be very worthwhile to tie the NAEP achievement level scores to something with great meaning to the public and the education community. The findings here offer hope that this may be possible, and that the NAEP is already providing important information that is currently unrecognized.

Richard G. Innes
October 2, 2018
70224.434@compuserve.com
Good afternoon-

I am writing as a private professional in educational measurement. My comments should be interpreted only as my own and should not be construed as representing my current employer. Disclosure: I have served in several capacities on NAEP projects over that last 25 years and was a member of the advisory panel convened by the Human Resources Research Organization (HumRRO) in July this year. I do not speak for other members of that panel. My comments follow.

First and foremost, I endorse the recommendations of the advisory panel and the revised ALs policy principles.

The seemingly small addition of the work "Content" to "Achievement Level Descriptors" provides important emphasis that these ALDs explicate the content area knowledge and skills targeted in NAEP assessments and required of examinees.

Principle 2.b emphases a modern concept in the validation of intended score interpretations and uses, that ALDs should be developed at the same time as assessment frameworks so that the assessment framework and all other design and development decisions and outcomes explicitly support the intended score interpretations and uses defined by the ALDs. This is the core of principled approaches to assessment design, development, and implementation and is fundamental to making rigorous validity arguments for the interpretation and use of NAEP assessment results.

Principle 2.c.i-iii and 2.d clarify nicely the idea that, while general public representation and input on NAEP ALs is crucial and ultimately democratic, representatives of the general public must be selected and trained so that they are well informed about the content area and students at the grade level under consideration in order to provide useful, well informed, and supported recommendations on AL cut scores.

Principle 3 in general, regarding validity evidence to support use of NAEP ALDs and data to interpret performance on NAEP assessments, is crucial to supporting such interpretations and uses.

However, principle 3.c may cause some confusion by stating that ALDs should reflect what students do know and can do, versus principle 2.b, which states that Content ALDs that guide AL setting should articulate what students should know and be able to do. Principle 3.c seems to be a reference to what are called Reporting ALDs in the standard setting literature. If so, to avoid confusion, why not use the term "Reporting ALDs" to distinguish them and their intended interpretations from Content ALDs.

Please let me know if you have questions about these comments.

Thank you,
Steve Ferrara, PhD
This review of NAEP achievement levels policies and practices is long overdue. These changes provide a more comprehensive policy while simplifying and clarifying the policy. A tremendous thanks to all who have contributed to the document.

Here are my suggestions. I hope they will be of help in finalizing this document.

Page 5, Principle 1, part a—I recommend adding this statement: "The content ALDs are the statement of the standard for each subject, grade, and level."

Page 5, Principle 1, part c—insert “know” so that the last sentence in that part reads: “what students in each achievement level know and can do.”

Page 5, Principle 2, part a—There should be a provision for coordination between the framework development group and the Assistant Director for Psychometrics who is responsible for the ALS process for developing the ALDs.

Page 6, Principle 2, part c. ii—Need to specify that each panel must meet these requirements.” The current statement is “This panel . . . .” Also, I think you need a minimum number of general public panelists specified. I think it is a bad idea to have fewer than 1 GP per table group, for example. As currently stated, the policy would allow as few as one or two GP panelists per grade panel.

Page 6, Principle 2, part c. iii—I would either omit the reference to shared (common) items or provide more information by saying that the item pool is also divided into two equivalent sets with a subset that is evaluated by both panel groups.

Page 6, Principle 2, part d—I feel pretty strongly that content facilitators should have experience with NAEP, specifically the NAEP framework for the subject area. That expertise is the way to assure that the content facilitators have a “claim” to authority over the panelists who may also be content experts. In addition, I feel strongly that the process facilitators must have some background and training in quantitative analysis. I have experienced the negative effects of having a process facilitator lacking that background and not being a successful process facilitator, despite being an excellent facilitator of groups in general.

Page 7, Principle 2, part e—The first sentence needs clarification—differ from what?

Page 7, Principle 2, part h—I think "testing out" is a little misleading. I think it should be made clear that any new features of the process should be tested prior to the pilot study, and any research questions must be evaluated prior to the pilot study.

Page 7, Principle 2, part i--TACSS has always been appointed by the contractor, with Board approval; and TACSS "reports" to the contractor—not the Board. This is important, and the current wording of the policy may be thought to suggest that the TACSS is appointed by the Board.
Page 7, Principle 2, part j, line 3—I suggest changing “at the time” to “by the time”

Page 8, Principle 2, part k, first sentence—I don't know if the Board still upholds this principle, but there was once a very strong notion that it was incorrect/inappropriate to refer to students in the achievement levels. Rather, one was to talk about the performance of students classified in the achievement levels and to say that the performance represents the knowledge and skills required by the ALDs.

Page 8, Principle 3, part b—The first sentence needs to be re-written to clarify that national results are available for all assessments and state and TUDA results for only some assessments. Also, I suggest omitting the last part of the second sentence because it does not really follow logically from the first part. I don't think that the prohibition on reporting necessarily means that the achievement levels do not apply to individual students or schools.

Page 8, Principle 3, part c, sentence 3—I think there will always be ALDs reported with results, so I suggest the sentence be edited to say: “In particular, the content ALDs for reporting results shall be written to incorporate...”.

Page 9, Principle 3, part f—At and above each cut score would result in overlapping regions! I think the levels need to be described as performance within the score ranges demarcated by each of the three cut scores with no upper limit to performance at the Advanced level.

Page 9, Principle 3, part g—I think this would be clearer and perhaps more accurate to say that NAEP Proficient may describe performance that is different from that in other assessments and that is not necessarily suggested by the common usage of "proficiency." (I think it is "proficiency" that gets things off-track!)

Page 9, Principle 4, part a—I note that this statement indicates that more than one NAEP assessment is to be involved in the review. I am pleased to see that!

Page 9, Principle 4, part b—Technically, this is fine, as written, but I feel some concern that this might imply that COSDAM, per se, will be expected to revise the ALDs. I think this statement should be clarified.

Page 9, Principle 4, part c—Again, I think it should be clear that this is to be coordinated with (actually, under the supervision of) the Assistant Director for Psychometrics because the ALDs must be developed with an understanding of how they are to be used in standard setting.

Page 9, Principle 4, part d—I think “revised” or “reset” sounds better than “redone.”
Page 10, Principle 5, part a—Suggest change to add some qualification about business representatives, such as: “business representatives in fields related to the content area.”

Page 10, Principle 5, part d—There is no indication of the qualifications required for TACSS members (for example, experience with NAEP, psychometric expertise, whether content knowledge is a consideration, representation of state assessment staff, representative of DAR contractor, etc.) nor the number of members for TACSS. There was also a requirement that at least one TACSS member have served on TACSS previously.

Page 11, Principle 5, part f—I am not sure this is needed since the Governing Board can ask anyone for input on the ALS process.

Page 11, Principle 6—general statement—Suggest modification to state: "... ensure that the final achievement level descriptions, cut scores, and exemplars recommended to the Governing Board for adoption comply with this policy."

Page 11, Principle 6, part b—This does not specify that COSDAM will and take action and make recommendations to the Board.

Page 11, Principle 6, part e, last line—Perhaps change “similar” to “relevant”.
Thank you for the opportunity to provide public comment on the review and possible revision of the NAGB policy on achievement levels.

As Commissioner of Education in Kentucky, I served on NAGB from 2011-2015 representing state superintendents.

I would encourage NAGB to maintain current cut scores and achievement levels to maintain the trend lines that have been established over the many years that NAEP has been administered. Given the many changes in education at the state and national levels over the past decade, it is critical that the "gold standard" of NAEP be maintained.

NAGB could certainly improve the communication about the achievement levels and provide clarity that proficient levels are set at the level of college and career ready expectations. Given the changing economy in our nation, it is critical that more high school graduates reach college and career readiness levels. NAEP is the "gold standard" to inform the public about our nation's performance and hold state assessments to the expectation that state achievement levels are set at rigorous levels that report on the preparation of students for college and career readiness.

Best wishes in your deliberation.

Terry Holliday
Response to request for comments on “Draft Policy Statement on Developing Student Achievement Levels for the National Assessment of Educational Progress,” posted Monday, September 10, 2018 by the National Assessment Governing Board, U.S. Department of Education in Notices of the Federal Register, Vol. 83, No. 175

These comments are authorized by the Steering Committee of the National Superintendents Roundtable, representing 75 school superintendents responsible for the education of some five million K-12 students, and 25 former superintendents.

Summary

The Roundtable appreciates the opportunity to comment on the draft policy statement. Although initially inclined to commend the National Assessment Governing Board (NAGB) for this effort to respond to the 2016 report of the National Academy of Sciences,¹ the Roundtable concludes that the proposed statement is unlikely to attain the goal defined in its first paragraph, namely to produce achievement levels for the National Assessment of Educational Progress (NAEP) that are “reasonable, useful, and informative to the public.”

The Roundtable supports high standards. While it believes NAEP misuses the term “Proficient,” it does not ask for a lowering of standards but for a clearer definition of what they mean in terms parents and the public can understand. In pursuit of that objective, we make the following observations: The original achievement levels were developed in a rushed process. Those levels have produced results have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent “wishful thinking” and defy “reason” and “common sense,” according to knowledgeable experts. The latest research linking NAEP’s benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP’s proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

We conclude with suggestions for rewriting the policy definitions and an appeal to extend the comment period from the currently contemplated 20 days to six months.

Rushed Initial process
Public Law 107-279, the Augustus F. Hawkins-Robert T. Stafford Elementary and Secondary School Improvement Amendments of 1988, created the National Assessment Governing Board and charged it with identifying appropriate achievement goals, which were established as “Basic,” “Proficient,” and “Advanced.”¹ An advisory panel appointed by NAGB in June 1990 to agree on the levels, and the proportion of students at each level who could be expected to answer questions correctly, reached agreement on these issues in November,² a remarkably brief period, given the importance and complexity of the issues involved. NAGB rejected the advice of experts to slow down the benchmarking process.³ NAGB board member Chester E. Finn, Jr. later dismissed the value of technical expertise: “I get fed up with technical experts [who] . . . take an adversarial stance toward some of the things that are most important the views of those operating NAEP, such as setting standards.”⁴

The results have confused educators, the public, and policymakers
Understandably the public, many commentators, and policymakers are confused by these benchmarks (achievement goals).

It is clear, as NCES states emphatically that “Proficient is not synonymous with grade-level performance.”⁵ According to Dr. Peggy Carr of NCES if the public wants to understand how many US students are performing on grade level it is more accurate to look at Basic scores, than Proficient results.⁶

Yet, on 2016, former television commentator Campbell Brown, newly leading an education advocacy organization called The 74 released a video on Slate arguing that “Two out of three eighth graders in this country cannot read or do math at grade level.” Asked for evidence of this claim she cited NAEP, apparently confusing the term Proficient with being at grade level.⁷

Major organizations such as Achieve Inc. and Students First misuse the term consistently in their publications, while Achieve and NCES issue analyses lamenting the fact that state definitions of proficiency fall far short of the NAEP Proficient benchmark.⁸

³ Vinovskis, M.A., op. cit.
Earlier this month, *Education Week* released its latest “Quality Counts” report. In this effort to “grade the states” on school quality, the report included state-by-state data on poverty gaps, high school graduation rates, and Advanced Placement experience. But the foundation of its achievement index is the proportion of students meeting the NAEP proficiency definitions in fourth- and eighth-grade reading and mathematics.

Oddly, the term Proficient as used by NAGB does not even mean proficient. As former NAGB officials wrote in 2001:

> Nor is performance at the Proficient level synonymous with ‘proficiency’ in the subject. That is, students who may be considered proficient in the subject, given the common usage of the term, might not satisfy the requirements for performance at the NAEP achievement level.

In short, although most citizens are understandably interested in performance at grade level, many confuse performance at grade level with NAEP’s proficiency benchmark. Understandably, journalists and many members of the public ignore the benchmark of Basic, interpreting it as an indication that large numbers of students are barely scraping by. This interpretation does great damage to public perceptions of public schools and perceptions of the ability and skills of the next generation of young Americans. To add to the confusion, officials associated with NAGB have openly acknowledged that NAEP’s definition of proficiency is unrelated to common usage of the term.

**Benchmarks represent “wishful thinking”**

The rushed process produced a scientific debate that lasts to this day.

NAGB hired a team of evaluators in 1990 to study the process involved in developing the three levels. A year later the evaluators were fired after their draft report concluded that the process “must be viewed as insufficiently tested and validated, politically dominated, and of questionable credibility.”

In 1993, the U.S. General Accounting Office labeled the standard-setting process as “procedurally flawed” producing results of “doubtful accuracy.”

In 1999, the National Academy of Sciences reported the achievement-level setting procedures were flawed: “difficult and confusing . . . internally inconsistent . . . validity evidence for the cut scores is lacking . . . and the process has produced unreasonable results.”

Shortly after *No Child Left Behind* was signed into law in 2001, Robert Linn, past president of the American Educational Association and of the National Council on Measurement in Education, characterized the three-level system as “misleading.”

---


10 Jack Jennings, *op. cit.*


Education, and former editor of the Journal of Educational Measurement called the “target of 100% proficient or above according to the NAEP standards is more like wishful thinking than a realistic possibility.”

In 2007, researchers concluded that fully a third of high school seniors who completed calculus, the best students with the best teachers in the country, could not clear the proficiency bar. Moreover, they added, fully 50 percent of those who scored “basic” in twelfth grade math had achieved a bachelor’s degree (a proportion comparing favorably with four-year degree rates at public universities).

The Buros Institute, named after the father of Mental Measurements Yearbook, criticized the lack of a validity framework for NAEP assessment scores in 2009 and recommending continuing “to explore achievement level methodologies.”

Fully 30 percent of 12th-graders who completed calculus were deemed to be less than proficient, said a Brookings Institution scholar in 2016, a figure that jumped to 69 percent for pre-calculus students and 92 percent for students who completed trigonometry and Algebra I. These data “defy reason” and “refute common sense,” he concluded.

Finally, the NAS study to which the proposed rule responds took note in 2016 of the “controversy and disagreement around the achievement levels, noting that Congress has insisted since 1994 that the achievement levels are to be used on a trial basis until on objective evaluation determined them to be “reasonable, reliable, valid, and informative to the public.”

In the Roundtable’s judgment, such an objective evaluation has yet to be completed and a determination that the achievement levels are “reasonable, reliable, valid, and informative to the public” has yet to be seen.

Linking studies conclude most students in most nations cannot clear “proficiency” bar

The Roundtable points also to research studies dating from 2007 to 2018 indicating NAEP’s proficiency bar is beyond the reach of most students in most nations. When Gary Phillips of the American Institutes of Research (and former Acting Commissioner of NCES) asked how students in other nations would perform if their international assessment results were expressed in terms of NAEP achievement levels, his results were sobering. The results demonstrated that just three nations (Singapore, the Republic of Korea, and Japan) would have a majority of their

---

students clear the NAEP bar in 8th-grade mathematics, while Singapore alone could meet that standard (more than 50% of students clearing the bar) in science.\textsuperscript{19}

Subsequently Hambleton, Sireci, and Smith (2007) and also Lim and Sireci (2017) reached conclusions similar to those of Phillips.

An analysis conducted by Gonulates and Harvey in 2017\textsuperscript{20} and reported by the Horace Mann League and the National Superintendents Roundtable in 2018\textsuperscript{21} applied Phillips’ methodology to international fourth-grade reading results as defined by the Progress in International Reading and Literacy Survey (PIRLS). PIRLS is conducted by the International Association for the Evaluation of Educational Achievement (IEA), the organization that administers the TIMSS mathematics and science assessments. Not a single nation participating in PIRLS can demonstrate that a majority of its students clear the NAEP reading proficiency bar in Grade 4. Singapore leads the way, with 39 percent of its students clearing the bar, followed by the Russian Federation (37 percent), Finland (36 percent), England (32 percent), and the United States (31 percent).

In light of these findings, it defies logic and common sense, as Dr. Loveless noted, to act on the belief that NAEP’s definition of proficiency is a reasonable and reliable guide to policy action.

**Proposed new policy definitions define anything less than “NAEP Proficient” as deficient**

Against that backdrop, it is surprising to see that NAGB proposes not to modify its policy definitions to clarify the term “Proficient” but to tinker at the margins with the definitions and add the word “NAEP” in front of each of the benchmarks. The justification offered in the Federal Register is to “better differentiate the NAEP achievement levels from other common uses of Basic, Proficient, and Advanced.” But the effect is to define anything less than NAEP Proficient as deficient.

The table on the following compares the existing policy definitions of the terms Basic, Proficient, and Advanced, with the policy definitions attached to the new terms “NAEP Basic,” “NAEP Proficient,” and “NAEP Advanced.” Like the existing policy definitions, the proposed definitions point inexorably to Proficient as the ideal to which states, the nation, and major demographic groups should, on average, aspire.

In the Roundtable’s judgment these minor modifications are in no way responsive to the major criticisms that have been leveled at the NAEP benchmarks over the years. To retreat behind the claim that the Proficient benchmark is an aspirational standard is deceptive and evasive. In light of the results of the international linking studies completed between 2007 and 2018, far from being aspirational, the Proficient benchmark is delusional. In Dr. Linn’s phrase, the proficiency benchmark represents wishful thinking.

\textsuperscript{19} It should be noted that no nation (or city or province) could demonstrate that its students met the NAEP definition of Advanced. (Cities and provinces, such as Chinese Taipei, Hong Kong are generally ignored for purposes of the analysis in this section.)

\textsuperscript{20} See Appendix C of How High the Bar?: “Applying NAEP Benchmarks to PIRLS results.”

<table>
<thead>
<tr>
<th>Benchmarks</th>
<th>Current Policy Definitions</th>
<th>Proposed Policy Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic</td>
<td>This level denotes partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at each grade.</td>
<td>This level denotes partial mastery of prerequisite knowledge and skills that are fundamental for performance at the NAEP Proficient level.</td>
</tr>
<tr>
<td>Proficient</td>
<td>This level represents solid academic performance for each grade assessed. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.</td>
<td>This level represents solid academic performance for each NAEP assessment. Students reaching this level have demonstrated competency over challenging subject matter, including subject matter knowledge, application of such knowledge to real world situations, and analytical skills appropriate to the subject matter.</td>
</tr>
<tr>
<td>Advanced</td>
<td>This higher level signifies superior performance.</td>
<td>This level signifies superior performance beyond NAEP Proficient.</td>
</tr>
</tbody>
</table>

Repairing the damage inflicted on public schools and their students by NAEP’s misleading terminology is all the more critical today as groups such as Achieve, Inc. openly advocate for NAEP’s definition of Proficient as the benchmark to be employed in state assessments under the *Every Student Succeeds Act*.

**Recommendations**

Against the backdrop of the discussion above, the Roundtable makes two recommendations:

1. Revise the proposed policy definitions so that they are more informative to the general public (one of the principal goals of the NAS analysis) We propose the following definitions:

<table>
<thead>
<tr>
<th>NAEP Basic</th>
<th>This level is roughly analogous to performance at grade level in each NAEP assessment. It denotes partial mastery of prerequisite knowledge and skills that are fundamental for performance at the NAEP Proficient level.</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAEP Proficient</td>
<td>This level represents extremely demanding academic performance, a level beyond the abilities of most students in most nations. Students reaching this level have demonstrated competency over challenging subject matter, including subject matter knowledge, application of such knowledge to real world situations, and analytical skills appropriate to the subject matter.</td>
</tr>
<tr>
<td>NAEP Advanced</td>
<td>This level signifies superior performance beyond NAEP Proficient.</td>
</tr>
</tbody>
</table>

2. Extend the comment period on the draft policy statement for six months, during which time NAGB should hold between four and six hearings around the nation to solicit the advice of educators, parents, and other stakeholders. George Santayana’s oft-quoted aphorism is usefully
born in mind here: Those who cannot remember the past are condemned to repeat it. The rushed process that produced the NAEP benchmarks in 1990 provides a cautionary tale for us today. NAEP has existed for nearly half a century. The benchmarks have been in use for over 25 years. Taking a few months to get it right on this occasion will be time well spent.

We commend NAGB for inviting comments and appreciate the opportunity to offer our insights. We stand ready to answer questions or help in any way with this process as it moves forward.

Sincerely,

James J. Harvey, Executive Director
(On behalf of Roundtable Steering Committee)

c.c. Peggy Carr, NCES

by email to NAEPALSpolicy@ed.gov
&
Peggy.Carr@ed.gov

And by overnight mail to:
NAEP Achievement-Level-Setting Program
National Assessment Governing Board
800 North Capitol Street, NW
Suite 825
Washington, D.C. 20002
I have been following the NAEP data for 30 years as a teacher, principal and superintendent. The data is always misrepresented in the media and to the public. I am very concerned that the term "proficiency" will again be changed, and misrepresented again!

I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Frank Ohnesorgen
Superintendent
Pond USD
29585 Pond Road
Wasco, CA 93280
To whom it may concern,

I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Sincerely,

Dr. Geoff Thomas
Superintendent
Madison 321
To Whom It May Concern,

I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Respectfully Submitted,

James H. Erinakes, II
Superintendent of Schools
Exeter-West Greenwich Regional School District
Rhode Island
Dear Ms. Carr,

I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Thanks for considering.
Jamey Harvey
CEO
Agilian.com
To Whom It May Concern,

I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are “reasonable, useful, and informative to the public,” the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Sincerely,
Jessica Essenter
Third Grade Teacher in Connecticut
Dear Department of Education:

I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Kim Fry
Superintendent
Rochester School District
RE: Response to draft policy statement on developing student achievement levels for the National Assessment of Educational Progress

To Whom It May Concern:

I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Thank you for considering my comments.

Kurt H. Hilyard
District Superintendent
Ocosta School District #172
2580 South Montesano Street
Westport, WA  98595
There is a lot of research out there that shows the current NAEP benchmarks are way above what are normal expectations. The research proves over and over again that no country can get most of their students to achieve at those benchmarks. The current benchmarks reflect more of a gifted standard which is only 10-15% of the population, not an expectation for the general populace. Also, the NAEP assessments are only given to a random sampling of students. They should never be used to rate or compare state or local education systems.

I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Laurie Vent
Superintendent  Upper Sandusky E. V. Schools
w: www.usevs.org e: laurie_v@usevs.org
Good afternoon:

My name is Matthew Montgomery and I am the superintendent of Revere Local Schools in Bath, Ohio.

I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Sincerely,

Matthew Montgomery

Matthew L. Montgomery
Superintendent
Revere Local Schools

Twitter: @SuptMontgomery
Revere Local Schools
I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

While I have cut and pasted the suggested comments, The current framework is not working and because all the scores are so low across the board, the reporting lacks value to students, parents, and staff.

Best regards,

Matthew G Scoggins
Rangely School District Re-4
I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP’s benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP’s proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

--

Michael J. Hoose, Superintendent
Dear Dr. Carr,

I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Sincerely,
Oren Pizmony-Levy

----

Oren Pizmony-Levy, PhD
Assistant Professor of International and Comparative Education
Department of International and Transcultural Studies
Teachers College, Columbia University
525 West 120th Street
370 Grace Dodge Hall
Box 55
New York, NY 10027
I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Richard Myles
Superintendent

Scottsbluff (NE) Public Schools
1722 First Avenue
Scottsbluff, NE 69361
Every coach sets the goal of being undefeated on the end of the season, but it just isn't possible for almost every team. I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Rob Busch
Superintendent/PK-6 Principal
Edgewood-Colesburg CSD
409 East Street – P.O. Box 125
Colesburg, Iowa 52035
I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Rob Terrill
District Assessment and Elementary Curriculum Director
Highland Local Schools
6506 SR 229
Marengo, Ohio 43334
To whom it may concern:

I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Sincerely,
--

Robert L. Cullison, Jr.
Superintendent
Prospect Mountain High School SAU #301
242 Suncook Valley Road
Alton, NH 03809
To Whom It May Concern:

I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. And in many of those nations, they educate only their best and brightest.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Shannon Peterson
Director

Main Campus
246 11th Ave SE
Forest Lake MN 55025 USA
651-464-0771

LILA Annex
121 11th Ave SE
Forest Lake MN 55025 USA
651-464-8989

Headwaters Campus
19850 Fenway Ave N
Forest Lake MN 55025 USA
651-464-8989
I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Thank you and I am able to give more information if someone would like.

Shawn
--
Shawn M Chabot
Assistant Superintendent/Chief Academic Officer
Lewiston Public Schools
I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Sincerely,

Steven D. Parker, Ed.D.
Superintendent
Lancaster County Public Schools
P.O. Box 2000
Kilmarnock, VA 22482
I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Sincerely,
Susette Bollard

--

Susette L. Bollard
Superintendent of Schools
Central Vermont Supervisory Union
Serving the Towns of Northfield, Orange, Washington, and Williamstown
To Whom It May Concern:

I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP’s proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Dr. Theresa R Rouse
Superintendent
Joliet Public Schools District 86
420 N. Raynor Ave.
Joliet, IL 60435
District Website: https://www.joliet86.org/
Superintendent's Blog: https://www.joliet86.org//superintendent-blog/
To whom it may concern:

I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I wish to note the following:

The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts.

The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar.

Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Tim O. Mains, Superintendent
Pine Bush Central School Dist.

twitter: @mainsPBsuper

*Knowing is not enough; we must apply.*
*Willing is not enough; we must do.*

~ Johann Wolfgang von Goethe
I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP’s benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP’s proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

--

Trina Evans
Liberty Middle School Principal
High School Mathematics Department
I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

--
Troy Parton, Superintendent
Munday CISD
PO Box 300
Munday, TX 76371
Home of the Moguls
Good morning-

I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts.

The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's "proficiency" bar. The proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days, now contemplated, to three months or more to give educators time to comment.

Thank you,

William M. Ward, Ed.D.
Superintendent of Schools
Pawling Central School District
Pawling, New York
I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP’s benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP’s proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Sincerely,

Brian Hummingbird
Principal
Warner High School
I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Chuck Benge
601 E. Madison
Fairfield, IA
Curriculum Director
I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Dr. Glen Fenter
Superintendent
Marion School District
200 Manor Street
Marion, AR 72364
Hello:

I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Danielle Bolduc, Gilford Elementary Principal
Gilford Elementary School
76 Belknap Mountain Road
Gilford, NH 03249
Dear National Assessment Governing Board,

My name is Ken Kunin, Superintendent of Schools in South Portland, Maine. We are a small suburban district of 3,000 students. I am writing regarding proposed changes to the NAEP achievement benchmarks.

I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Sincerely,
Ken Kunin
I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement. The current levels have produced results that have confused educators, policymakers, and families. The benchmarks represent "wishful thinking" and are inconsistent with children's development. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar.

I recommend that NAGB rewrite the policy definitions for its standards and extend the comment period from the 20 days now contemplated to three months or more to give educators time to provide informed and detailed comments.

Sincerely,

Iris C. Rotberg

Iris C. Rotberg, Ph.D.
Research Professor of Education Policy
Graduate School of Education and Human Development
The George Washington University
2134 G Street, NW
Washington, D.C. 20052
To whom it may concern,

As the superintendent of one of the largest school districts in Washington State, I am actively involved in legislation concerning both graduation requirements and student assessment. As such, I am very concerned with the direction that NAEP is presently taking.

I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Serving Evergreen’s Future

John Steach EdD
Superintendent
13501 NE 28th St.
Vancouver WA, 98682

Evergreen Public Schools

If we teach today's students as we taught yesterday's, we rob them of tomorrow.  John Dewey
The Department’s proposal to amend the policy definitions for NAEP’s benchmarks must be revised to make them more informative to the American public. Without such changes, most Americans will continue to believe that “Proficient” means performance at grade level.

Have a wonderful day!

MaryAnn Bragg
To Whom it May Concern:

I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research. I agree with the National Superintendent's Roundtable summary http://files.constantcontact.com/d6ed868c001/b63f4624-49ad-40a1-afd8-660913afe0cd.pdf and ask that feedback be considered.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Respectfully,

Steven T. Webb, Ed.D.
Superintendent
Vancouver Public Schools
2901 Falk Road
Vancouver, WA 98661

Connect with me on Twitter @SuptVPS
Dear Members of the National Assessment Governing Board,

I know that the national Assessment Governing Board is considering changes to the achievement benchmarks. I do not believe that the proposed changes for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," which is the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent wishful thinking and defy common sense, according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Sincerely,

Ms. Fern Lox
Assistant Principal
Chancellor Livingston Elementary School
Rhinebeck Central School District
Please revise the Department’s proposal to amend the policy definitions for NAEP’s benchmarks. These benchmarks should be changed to be more clear for parents and schools. Without such changes, most Americans will continue to believe that “Proficient” means performance at grade level.

Rob McEntarffer
To: U.S. Department of Education
Re: Jack McKay, Executive Director
Date: September 28, 2018

The Horace Mann League was founded a century ago to promote the benefits of public education for all in a
democratic republic. In support of our mission, the League was a major contributor to the development of a 2018
report, *How High the Bar? How Would Other Nations Perform if their Students were judged by Common Core or
NAEP Benchmarks?* In addition to the original research produced by that report, it contained a detailed literature
review of the many expert criticisms of NAEP's performance benchmarks, dating from their introduction in 1990 to
2018.

Against that background, I want to state clearly that the League's board does not believe the changes to the NAEP
achievement benchmarks proposed by the National Assessment Governing Board will produce achievement levels
that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed
policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end,
our board wants me to point out the following:
1. The original achievement levels were developed in a rushed process. Those levels have produced results that have
confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful
thinking" and defy "reason" and "common sense," according to knowledgeable experts.
2. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students
in most nations cannot clear NAEP's proficiency bar.
3. The proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history
of criticisms of the achievement levels and to the latest research.
4. Benchmarks or proficiency standards appears to be arbitrary and are set to have winner and losers. These
standards produce a zero sum game which seems to be contrary to the intent and mission of America's public
school. Using benchmarks and standards is folly when comparing America’s diverse population, limited social
services, and associated degrees of poverty.

Sincerely,

Jack McKay

Jack McKay, Ed.D., Executive Director
The Horace Mann League of the USA
Dear NAEP Achievement-Level-Setting Program,

As a former member of the National Assessment Governing Board, I am keenly interested in the improvement and credibility of the NAEP program.

I am writing to express my strong support for a complete rethinking of the NAEP "achievement levels." I urge the National Assessment Governing Board to abandon the achievement levels, because they are technically unsound and utterly confusing to the public and the media. They serve no purpose other than to mislead the public about the condition of American education.

The achievement levels were adopted in 1992 for political reasons: to make the schools look bad, to convey simplistically to the media and the public that "our schools are failing."

The public has never understood the levels. The media and prominent public figures regularly report that any proportion of students who score below "NAEP proficient" is failing, which is absurd. The two Common Core-aligned tests (PARCC and SBAC) adopted "NAEP Proficient" as their passing marks, and the majority of students in every state that use these tests have allegedly "failed," because the passing mark is out of reach, as it will always be.

The National Center for Educational Statistics (NCES) has stated clearly that "Proficient is not synonymous with grade level performance." Nonetheless, public figures like Michelle Rhee (who was chancellor of the DC public schools) and Campbell Brown (founder of the website "The 74") have publicly claimed that the proficiency standard of NAEP is the bar that ALL students should attain. They have publicly stated that American public education is a failure because there are many students who have not reached NAEP proficient.

In reality, there is only one state in the nation--Massachusetts--where as much a 50% of students have attained NAEP Proficient. No state has reached 100% proficient, and no state ever will.

When I served on NAGB for seven years, the board understood very well that proficient was a high bar, not a pass-fail mark. No member of the board or the staff expected that some day all students would attain "NAEP Proficient." Yet critics and newspaper consistently use NAEP proficient as an indicator that "all students" should one day reach. This misperception has been magnified by the No Child Left Behind Act, which declared in law that all students should be "proficient" by the year 2014.

Schools have been closed, and teachers and principals have been fired and lost their careers and their reputations because their students were not on track to reach an impossible goal.

As you well know, panels of technical experts over the years have warned that the achievement levels were not technically sound, and that in fact, they are "fatally flawed." They continue to be "fatally flawed." They cannot be fixed because they are in fact arbitrary and capricious. The standards and the process for setting them have been criticized by the General Accounting Office, the National Academy of Sciences, and expert psychometricians.
Whether using the Angoff Method or the Bookmarking Method or any other method, there is no way to set achievement levels that are sound, valid, reliable, and reasonable. If the public knew that the standards are set by laypersons using their "best judgment," they would understand that the standards are arbitrary. It is time to admit that the standard-setting method lacks any scientific validity.

When they were instituted in 1992, their alleged purpose was to make NAEP results comprehensible to the general public. They have had the opposite effect. They have utterly confused the public and presented a false picture of the condition and progress of American education.

As you know, when Congress approved the achievement levels in 1992, they were considered experimental. They have never been approved by Congress, because of the many critiques of their validity by respected authorities.

My strong recommendation is that the board acknowledge the fatally flawed nature of achievement levels. They should be abolished as a failed experiment.

NAGB should use scale scores as the only valid means of conveying accurate information about the results of NAEP assessments.

Thank you for your consideration,

Diane Ravitch
NAGB, 1997-2004
Ph.D.
New York University
Dear NAEP Policy Board:
I strongly support the views of Diane Ravitch expressed in her recent letter to you, advocating that "... the board acknowledge the fatally flawed nature of achievement levels. They should be abolished as a failed experiment. NAGB should use scale scores as the only valid means of conveying accurate information about the results of NAEP assessments."

Sincerely,

Dan Drmacich, Coordinator; Rochester Coalition for Public Education
September 19, 2018

To whom it may concern,

I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Respectfully,

Bobby Waitman
I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Thanks for your consideration.

--
Carl Bruner, Ed.D.
Superintendent
Mount Vernon School District No. 320
www.mountvernonschools.org
Dear National Assessment Governing Board:

I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Sincerely,

Knoxville Panthers
Knoxville Community School District
309 W Main Street, Knoxville, IA 50138
o: (641) 842-6551 m: (641) 751-5891
e: cassi.pearson@kcsd.k12.ia.us
Thursday, September 20, 2018

To Whom It May Concern:

I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Sincerely,

Derrick Meador, Superintendent
Jennings Public Schools
475 North Oak Street
Jennings, OK 74038
918-757-2536
The Department’s proposal to amend the policy definitions for NAEP’s benchmarks must be revised to make them more informative to the American public. Without such changes, most Americans will continue to believe that “Proficient” means performance at grade level. I support the findings of “How High the Bar,” and I implore you to study that report further.

Martha Bruckner, Ph. D.
Executive Director, MOEC Collective Impact Initiative
Metropolitan Omaha Education Consortium
UNO: Barbara Weitz Community Engagement Center, 223 B
6001 Dodge Street, Omaha, NE 68182
I do not believe the changes proposed by the National Assessment Governing Board to the achievement benchmarks for the National Assessment of Educational Progress will produce achievement levels that are "reasonable, useful, and informative to the public," the goal defined in the first paragraph of the proposed policy statement.

These standards and benchmarks must be expressed in terms that parents and the public can understand. To that end, I point out the following: The original achievement levels were developed in a rushed process. Those levels have produced results that have confused educators, citizens, and policymakers. Instead of being reasonable, the benchmarks represent "wishful thinking" and defy "reason" and "common sense," according to knowledgeable experts. The latest research linking NAEP's benchmarks to international assessments reveals that the majority of students in most nations cannot clear NAEP's proficiency bar. Finally, the proposed modifications to the policy definitions for NAEP achievement levels are unresponsive to the history of criticisms of the achievement levels and to the latest research.

NAGB should rewrite the policy definitions for its standards and extend the comment period from the completely inadequate 20 days now contemplated to three months or more to give educators time to comment.

Jeff Morrison
ETHICS PRIMER

FOR

THE NATIONAL ASSESSMENT GOVERNING BOARD

October 2018
Ethics Division
Office of the General Counsel
U.S. Department of Education
EXECUTIVE SUMMARY ............................................................................................................ 3

ETHICS LAWS AND RULES APPLICABLE TO SGES ..................................................................... 6

I. INTRODUCTION ................................................................................................................... 6

II. YOUR STATUS AS A SPECIAL GOVERNMENT EMPLOYEE ........................................... 6
   A. What is a “special Government employee”? ............................................................... 6
   B. Do the ethics restrictions apply when I am not working for NAGB? ......................... 6

III. CONFLICTS OF INTEREST .............................................................................................. 6
   A. What criminal conflict of interest statutes apply to SGEs? ............................................ 6
   B. What financial conflicts of interest may arise for SGEs under section 208? ................. 7
   C. How do I resolve a conflict of interest? ..................................................................... 9
   D. What restrictions apply to my representation of third parties under sections 203 and 205? .......................................................... 10

IV. POST-EMPLOYMENT ..................................................................................................... 10

V. STANDARDS OF ETHICAL CONDUCT AND OTHER ETHICS RULES ....................... 10
   A. What restrictions apply if I want to engage in fundraising? ...................................... 11
   B. What restrictions are there on my acceptance of gifts? ............................................. 11
   C. What restrictions apply if I want to “lobby” Congress? ............................................ 12
   D. What does “misuse of position” mean? ................................................................. 12
   E. May I keep my day job and still serve on NAGB? ..................................................... 12
   F. Are there any restrictions on my political activities? ................................................. 12
   G. What restrictions do I face if I want to teach, speak, or write on matters that are related to the duties I perform for NAGB? ......................... 13
   H. What restrictions apply if my government duties involve the awarding of contracts? 13
   I. What restrictions apply to my interaction with foreign entities? ................................. 14
   J. What do I do if I am called to be an expert witness? .............................................. 14
   K. May I keep and use frequent flyer miles that I earn when I am on official NAGB travel? ........................................................................... 15

VI. CONCLUSION .................................................................................................................. 15
EXECUTIVE SUMMARY

Now that you are a member of the National Assessment Governing Board (“NAGB”) you need to know what ethics laws and rules apply to you. The following is a very brief summary of these rules. For a more detailed discussion of how these rules apply to you, please refer to the attached summary entitled “Ethics Laws and Rules Applicable to SGEs.”

Your Status as a Special Government Employee

You are considered an SGE and not a regular federal employee because NAGB anticipates that you will be serving the federal government through your position for only 130 days or less during any period of 365 consecutive days. Whether or not you are paid by the Board for your service is irrelevant. This summary discusses how the ethics rules apply to SGEs.

Criminal Statutes Apply to Your Activities

Some of the ethics laws that apply to you carry criminal penalties. Below is a brief summary of the most important of these laws.

- The chief conflict of interest law bars you from participating personally and substantially in your capacity as a member of NAGB in any particular matter before the federal government that has a direct and predictable effect on your own financial interests or the financial interests of others with whom you have certain relationships. See 18 U.S.C. Section 208.

- If you find yourself with a financial conflict of interest, you have four options: (1) disqualify yourself (you don’t participate in any way in the matter); (2) resign from the outside entity that is the basis for the conflict; (3) sell or divest the stock or other financial interest that is the basis for the conflict; or (4) request and obtain a statutory waiver.1

- Two other laws prohibit you from representing a third party, with or without compensation, before any court or agency in connection with any particular matter involving specific parties in which the United States is a party or has a direct and substantial interest and in which you have participated personally and substantially as an SGE. In addition, if you serve the federal government for more than 60 days during the immediately preceding period of 365 consecutive days, these restrictions apply to any matter that is pending with NAGB. But remember that these restrictions do not apply to particular matters of general applicability, such as broadly applicable policies, rulemaking proceedings or legislation, that do not involve specific parties. See 18 U.S.C. Sections 203 and 205.

1 In rare circumstances, with the concurrence of the U.S. Office of Government Ethics, you may obtain a waiver of the conflict of interest.
• Another criminal law limits some of your activities after your service on NAGB ends. This law prohibits you from representing others in connection with the same particular matter involving specific parties in which you participated personally and substantially during your service to NAGB. This prohibition lasts for your lifetime. See 18 U.S.C. Section 207.

Standards of Ethical Conduct for Employees of the Executive Branch

The Standards of Ethical Conduct for Employees of the Executive Branch (Standards), 5 C.F.R. Part 2635, are regulations that apply both to regular federal government employees and to SGEs. However, a few exceptions exist in the Standards in recognition of the fact that SGEs are working for the government only in a very limited way. A brief synopsis of some these rules and their exceptions follow.

• **Fundraising:** You may not use your official title, position and authority to engage in fundraising.

• **Gifts:** You may not accept gifts from a “prohibited source” or offered to you because of your official position on NAGB. A prohibited source includes any person: seeking official action from NAGB; doing or seeking to do business with NAGB; conducting activities regulated by NAGB; or having interests that may be substantially affected by your official duties. There are many exceptions to this rule that are discussed in more detail in the accompanying memorandum.

• **Lobbying:** In your role as a member of NAGB, you may not urge others to contact Congress or a state legislature to urge the passage or defeat of legislation. Additional restrictions exist regarding lobbying. You should contact Department of Education’s Ethics Division before engaging in any type of lobbying.

• **Misuse of Position:** You may not use your position on NAGB or nonpublic information gained through your service on NAGB to seek advantage for yourself or others. In addition, you may not use your NAGB title in a manner that makes it appear that NAGB is sanctioning your views, products, services or personal enterprises.

• **Political Activities:** You may not engage in political activity when you are on duty or in a federal government building or car, and you may never use your official title as a member of NAGB in connection with political activities.

• **Teaching, Speaking and Writing:** You may not receive compensation for teaching, speaking or writing if: (1) the invitation was offered to you because of your position on NAGB; (2) the information conveyed by you draws substantially on nonpublic information that you obtained by working on NAGB; (3) the invitation was extended to you by an organization or person who has interests that may be substantially affected by your performance on NAGB; or (4) the subject of your work deals in a significant way
with a matter involving specific parties that you worked on while on NAGB. Again, there are some exceptions to this rule that are outlined in more detail in the accompanying memorandum.

**Required Filing of a Financial Disclosure Report By SGEs**

As a member of the NAGB, you are required to file a confidential financial disclosure report (also referred to as a “450” Report) when you are first appointed, and annually thereafter if you are reappointed. The purpose of the financial disclosure form is to protect you from inadvertently violating any of the criminal conflict of interest statutes and so that NAGB can know that your advice is free from any real or perceived conflicts of interest.

Please do not rely solely on this “Executive Summary” before undertaking your duties. There are many subtle nuances that are not discussed in this summary that may apply to your specific situation. The attached expanded summary provides additional detail that will help you better understand the ethics rules. Please feel free to call or e-mail Marcella Goodridge in the Ethics Division of the Office of the General Counsel at the U.S. Department of Education at (202) 401-8309, or Marcella.Keiller@ed.gov, for answers to any specific ethics questions that may arise in the course of your service on NAGB.
ETHICS LAWS AND RULES APPLICABLE TO SGES

I. INTRODUCTION

Although the ethics rules are numerous and detailed, a single, simple principle underlies these rules: *You should never use your public office for private gain, either for yourself, or for any third party.* In addition, you must refrain not only from engaging in any activity that violates the ethics rules, but you must also refrain from any activity that creates the appearance of a violation of any of these rules. The summary below is designed to help you avoid violating any ethics rules covering your activities as a member of NAGB.

II. YOUR STATUS AS A SPECIAL GOVERNMENT EMPLOYEE

A. What is a “special Government employee”?

Because you have been appointed to be a member of the NAGB and you are expected to perform your duties for not more than 130 days during the 365 days subsequent to the date of your appointment, you are, by law, a “special Government employee” (SGE). As an SGE, you are a federal government employee. This means that upon appointment, you assume the responsibilities, obligations, and restrictions that are part of public service. Because SGEs are not full-time employees, several of these restrictions apply only in limited circumstances.

B. Do the ethics restrictions apply when I am not working for NAGB?

Yes, any restrictions concerning your private activities (representational services, expert witness activities, etc.) apply equally on days when you serve the federal government through your position on NAGB and on days when you do not, except with respect to political activity. If you have not provided any services for the federal government for some time, but have not received a termination date for your appointment, you must seek a formal resolution of the matter before engaging in conduct prohibited by the ethics rules.

III. CONFLICTS OF INTEREST

A. What criminal conflict of interest statutes apply to SGEs?

While you are employed as an SGE, you need to pay particular attention to four criminal conflict of interest laws found in Chapter 11, Title 18 of the United States Code: 18 U.S.C. Sections 203, 205, 207 and 208. These criminal laws include some special provisions for the treatment of SGEs. A discussion of these laws and certain related requirements found in other laws and regulations follows.
B. What financial conflicts of interest may arise for SGEs under section 208?

Section 208 prohibits you from participating personally and substantially in any particular matter that has a direct and predictable effect on your financial interests, including certain interests of others that are imputed to you under the statute. This means that you may not work on NAGB matters if you have certain connections – through the ownership of stock, through employment, or by virtue of other circumstances – with an organization that has a financial interest in the matter. For example, you may not work at all on a contract competition if you own stock valued at a certain amount in a company competing for the contract. You may not participate in a discussion of whether to modify an existing contract with a company if you work for that company. And, you may not assist in the development of a scope of work for a contract competition if you know that an organization on which you serve on the Board of Directors plans to compete for that contract.

In addition to your own personal financial interests, the financial interests of the following persons or organizations are imputed to you and also disqualify you from participating in a particular matter:

   (1) your spouse;
   (2) your minor child;
   (3) your general partner;
   (4) an organization for which you serve as an officer, director, trustee, general partner or employee; and
   (5) any prospective employer.

Example 1  You are on the governing board of ABC, a nonprofit organization. ABC’s financial interests are imputed to you under the statute. This means that for the purpose of determining whether you have a conflict of interest, ABC’s financial interests are treated as if they were your own. Accordingly, you may not participate in any NAGB matter in which ABC has a financial interest. Similarly, if you were in the process of discussing employment with ABC, you would be barred from participating in any NAGB matter affecting the financial interests of ABC.

Example 2 You are on the governing board of ABC (or employed by ABC, own stock in ABC, seeking employment with ABC, etc). You are asked to participate in the process of reviewing and scoring contract proposals for a contract competition for a NAGB project. Fifteen organizations have submitted a bid. When you open the proposal from one organization, you note that ABC’s name is one of the organizations that has submitted a bid. Or, perhaps ABC is listed as a subcontractor in one of the proposals. This contract competition is a “particular
NOTE: Apart from the criminal conflicts of interest statutes discussed above, a regulation also exists that prohibits you from participating in a matter involving specific parties if a reasonable person would question your impartiality.

You must recuse yourself from a matter as soon as you realize that you have a conflict. If, for example, you notice that you have a conflict when you are in the middle of reviewing contract proposals, you put the proposal back in its envelope and call up an NAGB staff member and let that person know that you think that you are disqualified from working on the competition. If there is any question, you should contact the U.S. Department of Education Office of the General Counsel’s Ethics Division for guidance. Once you have determined that you may not work on this matter, send the proposal back to NAGB staff.

You are permitted to participate in a particular matter affecting one campus of a multi-campus institution of higher education, where the disqualifying interest arises from your employment with a separate campus of the same institution, provided that you have no multi-campus responsibilities at the institution. If you are employed with a large university with multiple campuses and you do not have any multi-campus responsibilities, you may participate in official matters—such as grants, contracts, applications, and other particular matters—that affect the financial interests of another campus in the same university system where you are employed. Below are some examples of how section 208 may apply to your activities.

Example 3 You are employed as a professor at the University of California-Berkeley. NAGB is planning to evaluate the impact of computer-based testing on students with disabilities and English language learners. UC-Berkeley’s science and technology department has submitted a bid. NAGB’s actions will have a direct and predictable effect on the university’s financial interest. Therefore, you may not participate in any way on this matter.

Example 4 You are employed as a researcher at the University of California-Berkeley. NAGB is planning to evaluate the impact of computer-based testing on students with disabilities and English language learners. The University of California-Los Angeles (UCLA) has submitted a bid to be the contractor for NAGB’s evaluation. You may participate in this matter because it will not have a direct and predictable effect on either your financial interests or UC-Berkeley’s.
C. How do I resolve a conflict of interest?

1. Disqualification

A common method of resolving a conflict of interest is to disqualify yourself from participating in the matter.

*Example 5* You are serving on NAGB’s Ad Hoc Committee that will examine issues related to computer-based testing for students with disabilities and English language learners, including developing a study of computer-based testing methodologies. The Request for Proposals has been disseminated. One of the bids submitted is from ABC Corporation (ABC). You own $20,000 worth of stock in ABC. You must advise the U.S. Department of Education Office of the General Counsel’s Ethics Division that you own stock in ABC and you will not be able to participate in any way in the entire contract competition. If ABC is awarded the contract, you will also need to disqualify yourself from the entire matter.

2. Divestiture

Divestiture of a disqualifying interest (usually through the sale of stock) is another remedy available to avoid a potential violation of section 208. SGEs are not eligible for a Certificate of Divestiture (CD). A CD is a tax benefit that allows the deferral or nonrecognition of capital gain where an employee divests a financial interest in order to comply with conflict of interest requirements. Unfortunately, Congress specifically excluded SGEs from eligibility to receive CDs. 26 U.S.C. § 1043(b)(1)(A).

3. Resignation

On some very rare occasions when none of the aforementioned options are available or feasible, an SGE may need to resign from participating in an outside activity with an entity if his or her official activities as an SGE have a direct and predictable effect on the financial interest of that entity creating an irreconcilable conflict.

4. Waiver or Authorization

Another remedy to avoid a conflicting financial interest is to request and obtain a statutory waiver by contacting the Department of Education’s Ethics Division (an authorization is similar to a waiver, but only applies to non-statutory conflicts of interest - what are often referred to as “appearances of a conflict”). You may be granted a waiver only if your financial interest is not so substantial as to be deemed to be likely to affect the integrity of your services.

*Example 6* In the scenario described in Examples 1 and 2 above, you are granted a waiver permitting you to participate in a general policy matter that affects ABC’s financial interests as

Any waiver or authorizations that you receive will be limited. It is very important that you read it carefully, as it will often contain detailed information about the types of matters from which you remain disqualified, despite the waiver or authorization.
long as the matter affects all similarly situated entities in the same manner. But you would remain disqualified from participating in a matter that specifically involves ABC, which in this case means the entire contract competition.

D. What restrictions apply to my representation of third parties under sections 203 and 205?

With regard to particular matters in which you have participated personally and substantially while serving NAGB, you are prohibited from representing a third party on those particular matters, with or without compensation, before any court or agency, when the United States is a party or has a direct and substantial interest in the matter. See 18 U.S.C. Sections 203 and 205.

In addition, if you serve the federal government for more than 60 days during the immediately preceding period of 365 consecutive days, you are prohibited from representing a third party on any matter involving specific parties pending before NAGB, even if your work at NAGB did not involve these matters. These restrictions do not apply to particular matters of general applicability, such as broadly applicable policies, rulemaking procedures or legislation that does not involve specific parties.

IV. POST-EMPLOYMENT

After your appointment terminates at NAGB, you need to pay particular attention to one more criminal statute that subjects you to restrictions regarding certain matters that you may have worked on as a member of NAGB. Pursuant to 18 U.S.C. Section 207, you may never represent any third party, other than in the performance of your official government duties, in connection with the same particular matter involving specific parties in which you participated personally and substantially as a member of NAGB. This is a lifetime prohibition. For example, if you participated in a NAGB discussion concerning a contract to State University, you may never represent State University with respect to that same contract before any official of the Executive Branch of the federal government and you may never represent State University with respect to that contract in any federal court.

Further, if you serve on NAGB more than sixty days and are compensated above a certain level, you may be subject to a one-year “cooling-off” period during which you would be barred from representing before NAGB certain third parties in connection with any matter. There are some exceptions to this law as well, and you should contact the Department of Education’s Ethics Division for guidance.

V. STANDARDS OF ETHICAL CONDUCT AND OTHER ETHICS RULES

The Standards of Ethical Conduct for Employees of the Executive Branch (Standards), 5 C.F.R. Part 2635, are regulations that apply both to regular federal government employees and to SGEs. Although you are treated generally the same as regular employees under the Standards, a few
exceptions do exist for SGEs in recognition of the fact that SGEs are working for the government only in a very limited way. In addition, there are other rules that govern your conduct as an SGE, including the Hatch Act, anti-lobbying rules, the Federal Acquisition Regulation, and rules about accepting gifts and compensation from foreign governments. A brief synopsis of some of these rules follows.

A. What restrictions apply if I want to engage in fundraising?

You may not use your NAGB title, position or authority to solicit funds for any organization. In addition, you may not personally solicit funds or other support from persons whose interests may be affected substantially by the performance or nonperformance of your official duties.

B. What restrictions are there on my acceptance of gifts?

You are prohibited from accepting gifts (almost anything of monetary value) from a “prohibited source” or gifts given because of your official position as a member of NAGB, unless a specific exception applies. The definition of “prohibited source” includes any person:

- seeking official action from NAGB;
- doing or seeking to do business with NAGB; or
- having interests that may be substantially affected by your official duties at NAGB.

The definition also includes organizations the majority of whose members fall within any of these categories. You may accept various benefits resulting from your outside business or employment activities, if a reasonable person would conclude that such benefits are not offered or enhanced because of your official position. The most commonly applicable exceptions to the gift rule allow you to accept:

- Modest items of food other than a meal, such as coffee, soft drinks, or donuts;
- Most plaques, certificates and trophies;
- Discounts available to all Government employees;
- Anything for which you pay market value;
- Gifts valued at $20 or less per occasion, totaling no more than $50 in a calendar year from any one source;
- Gifts clearly motivated by friendship or family relationship;
- Gifts resulting from your outside business activities, including those of your spouse; and
- Free attendance or meal which is provided by:

1. The sponsor of the event for the day on which you are speaking or you are otherwise expected to present information on behalf of the Board at the event, or for a widely-attended gathering of mutual interest to a number of parties when the necessary determination of agency interest has been made; or

2. someone other than the sponsor of a widely-attended gathering of mutual interest to a number of parties when more than 100 people are expected to attend, the
aggregate value of the gift is under $335, and the necessary determination of agency interest has been made.

C. What restrictions apply if I want to “lobby” Congress?

NAGB and its members are permitted to communicate directly with Congress in their official capacity on matters that are related to legislation or appropriations deemed necessary to conduct NAGB’s “public business” (i.e., the NAGB’s statutory functions and responsibilities). However, the Anti-Lobbying Act, 18 U.S.C. Section 1913, prohibits you, in your official capacity at NAGB, from engaging in “grass-roots lobbying” (i.e., directly or indirectly suggesting or requesting that others contact Congress or a state legislature to urge the passage or defeat of proposed or pending legislation), even if it is related to the NAGB’s public business. The Anti-Lobbying Act also requires that any permissible direct communications with Congress in your official capacity at NAGB be made only through official channels.

None of these restrictions prohibit you from lobbying members of Congress or state legislatures, or urging others to do so, on your own time in your personal capacity. If you lobby Congress or state legislatures in your personal capacity, and the issue is related to NAGB’s business, you should make it clear that you are not representing NAGB and not acting in your official capacity as a member. Also, please note that when you are lobbying as a private citizen, you are not permitted to use government resources or equipment (including, but not limited to, computers, telephones, fax machines, copy machines, stationery), or seek assistance from NAGB staff.

D. What does “misuse of position” mean?

You may not use your position on NAGB to seek advantage for yourself or others. You also may not use nonpublic information gained through your service at NAGB to seek advantage for yourself or others. Finally, you may not use your NAGB title in a manner that makes it appear that the NAGB is sanctioning your views, products, services or personal enterprises. Of course, you may list your membership on NAGB on your curriculum vitae, but you may never use your status as an NAGB member to advertise or promote your personal activities. Please seek advice from the Department of Education Office of the General Counsel’s Ethics Division if you have any questions in this area.

E. May I keep my day job and still serve on NAGB?

Yes, you may continue to collect your regular salary from an outside employer for days on which you are providing services to the federal government (whether your federal government service is paid or unpaid). However, if you have another consultant or advisory position with NAGB or any other federal department or agency, you may not receive per diem or salary from NAGB for the same day for services performed for the two positions.

F. Are there any restrictions on my political activities?

You may not engage in any political activities while you are on duty (i.e., performing
government services) or when you are in a government building or vehicle. Although you are not subject to any restrictions on your political activities when you are not performing government services, you may never use your official title as a member of NAGB in connection with any political activities.

G. What restrictions do I face if I want to teach, speak, or write on matters that are related to the duties I perform for NAGB?

You may not receive compensation for teaching, speaking, or writing if:

- the activity is performed as part of your official duties (e.g., a speech on behalf of NAGB);
- the invitation to engage in the activity was extended primarily because of your official position at NAGB, rather than expertise in the subject matter;
- the invitation or offer of compensation was extended to you by someone with interests that may be affected substantially by your duties;
- the information conveyed through the activity draws substantially on nonpublic information obtained through your service at NAGB; or
- the activity deals, in significant part, with a matter involving specific parties to which you are currently assigned or had been assigned during your current NAGB appointment.

Notwithstanding the restrictions in bold type you may accept compensation for teaching a course requiring multiple presentations offered as part of: (a) the regularly established curriculum of various specified types of educational institutions; or (b) educational or training programs sponsored and funded by federal, State, or local government. However, if you teach at an educational institution, you must not participate in any NAGB matters that involve that institution.

H. What restrictions apply if my government duties involve the awarding of contracts?

If you are involved in the awarding of any contracts, please seek advice from the Ethics Division. There are special provisions that cover your involvement in the awarding of contracts. For example, you may not accept compensation as an employee, officer, director, or consultant of a contractor within the one-year period after leaving Government service where you participated in certain procurement matters pertaining to that contractor. In addition, if you disclose certain information pertaining to Federal procurements that you obtained during your service on a committee, you may face sanctions, including criminal penalties.
I. What restrictions apply to my interaction with foreign entities?

The emoluments clause of the U.S. Constitution prohibits you from receiving any emolument, office or title of any kind from a foreign government, including political subdivisions of a foreign government. An emolument is compensation received by virtue of holding an office or having employment with a foreign government and includes, for example, salary, honoraria, transportation, per diem allowances, household goods, shipment costs, and housing allowances. This clause has been interpreted to be broader than the traditional notion of employment and includes, for example, income received through a partnership when an identifiable portion of the partnership draw can be attributed to the partnership’s fees from such foreign government. This provision has particular relevance to positions with foreign universities that are government-operated, as opposed to private institutions. United States Constitution, art. I § 9, cl. 8. There are also statutory provisions restricting acceptance of gifts from foreign governments. 5 U.S.C. § 7342. You should seek advice from the Ethics Division regarding the details about these restrictions. Additionally, a criminal statute bars employment or consultation with a foreign entity for the purpose of providing foreign agent representation or lobbying. 18 U.S.C. § 219.

The ban on participating in foreign agent activities covered by the Foreign Agents Registration Act (FARA) prohibits representation of foreign governments or foreign political parties before the United States Government, as well as a number of other activities conducted within the United States on behalf of such entities. There are certain FARA exceptions related to trade or commerce, legal representation, humanitarian fundraising, and religious, scholastic, or scientific pursuits. The Lobbying Disclosure Act of 1995 requires certain covered Federal officials who serve as agents of foreign principals (other than foreign governments or foreign political parties) to register if they work on behalf of foreign corporations, associations, or other organizations.

Finally, certain restrictions apply after your position with NAGB terminates. Specifically, 18 U.S.C. § 207 includes restrictions on former employees who participated in trade or treaty negotiations on behalf of the United States (18 U.S.C. § 207(b)) and on former senior employees who wish to represent, or aid or advise in the representation of, a foreign entity with the intent to influence a decision of a Federal employee or agency (18 U.S.C. § 207(f)).

J. What do I do if I am called to be an expert witness?

Government employees generally may not participate as an expert witness, with or without compensation, other than on behalf of the United States, in any proceeding before a federal court or agency in which the United States is a party or has a direct and substantial interest. This restriction applies to most SGEs only if the SGE actually participated officially in the same proceeding or in the particular matter that is the subject of the proceeding. If you are appointed by the President, serve on a commission established by statute, or serve (or are expected to serve) for more than 60 days in a period of 365 days, the restriction on expert service also applies to any proceeding in which NAGB is a party or has a direct and substantial interest.
K. May I keep and use frequent flyer miles that I earn when I am on official NAGB travel?

Yes, you may use frequent flyer miles or other airline awards or promotions accumulated on official NAGB travel for your own personal use.

VI. CONCLUSION

We understand that these laws are complex and may not be intuitive. Again, we caution you that this summary is merely an introduction to the ethics laws and rules that apply to you. You should always feel free to contact the Department of Education Office of the General Counsel’s Ethics Division with any questions or concerns.

Marcella Goodridge Keiller, Attorney
U.S. Department of Education
Office of the General Counsel
400 Maryland Avenue, S.W., Room 6E237
Washington, D.C. 20202-2110
(202) 401-8309
(202) 260-5104 (fax)

Marcella.Keiller@ed.gov
National Assessment Governing Board
Nominations Committee

November 17, 2018
7:30 – 8:20 am

AGENDA

7:30 – 7:35 am  Welcome, Introductions, and Agenda Overview
                Fielding Rolston, Chair

7:35 – 7:45 am  Update and Discussion of Nominations Received for 2019
                Fielding Rolston

7:45 – 8:05 am  Final Review of the Procedures Manual
                Fielding Rolston

8:05 – 8:15 am  Review of Candidate Rating System
                Lisa Stooksberry

8:15 – 8:20 am  Next steps
                Fielding Rolston

8:20 am        Adjourn
Report of Recommendations Offered by the
Ad Hoc Committee on Measures of Postsecondary Preparedness

November 17, 2018

In November 2016, the Governing Board unanimously approved its Strategic Vision to focus the Board’s work through the year 2020. Included within that vision is the priority to, “Develop new approaches to measure the complex skills required for transition to postsecondary education and career.” In August 2017, the Governing Board established the Ad Hoc Committee on Measures of Postsecondary Preparedness (committee) to review existing research, collect expert testimony, and prepare recommendations for the Governing Board’s consideration. Over the course of a year, between November 2017 and November 2018, the committee met quarterly and commissioned several research papers and convened five expert panel meetings in various regions of the country. The summaries of these are included in the draft appendix on the following pages.

On Saturday, November 17, 2018 the committee will report their final recommendation to the full Board. This report will complete the committee’s charge. No Board action is required at this meeting. The recommendations report is not included in this document; it will be provided to Board members separately in advance of the meeting.
Table of Appendices

Appendix A. Charge to the Committee ................................................................. A-2
Appendix B. Exploratory Approach Overview ................................................... B-4
Appendix C. Summary Answers for Research Questions ................................. C-6
Appendix D. Expert Panel: Industry ................................................................. D-15
Appendix E. Expert Panel: Higher Education .................................................. E-26
Appendix F. Expert Panel: Futurists ................................................................. F-39
Appendix G. Expert Panel: State Assessment Directors ............................... G-85
Appendix H. Expert Panel: Youth ................................................................. H-93
Appendix I. Literature Review: Work of the Future ...................................... I-95
Appendix J. Literature Review: Skills of the Future ........................................ J-114
Appendix K. Literature Review: State Indicators of College and Career Preparedness .... K-142
Appendix A. Charge to the Committee
Resolution: The Executive Committee’s Charge to the Ad Hoc Committee on Measures of Postsecondary Preparedness

Whereas, on November 18, 2016, the National Assessment Governing Board unanimously approved the Strategic Vision to guide its work through the year 2020; and

Whereas, the Strategic Vision established a Board priority (SV#10) to “Develop new approaches to measure the complex skills required for transition to postsecondary education and career”; and

Whereas, on August 3, 2017, the Governing Board Chair created the Ad Hoc Committee on Measures of Postsecondary Preparedness to pursue this priority; and

Whereas, the Governing Board Chair tasked the Executive Committee to establish the charge to guide the Ad Hoc Committee on Measures of Postsecondary Preparedness;

Therefore, the Executive Committee resolves that:

1. The Ad Hoc Committee on Measures of Postsecondary Preparedness shall review existing research, collect expert testimony, and prepare recommendations for the Governing Board’s consideration to achieve Strategic Vision priority #10.

2. While the current legislation guiding the National Assessment of Educational Progress (P.L. 107-279) should provide parameters for the approaches to accomplish this priority, the Ad Hoc Committee on Measures of Postsecondary Preparedness may consider options that could require amendments to current legislation.

3. The Ad Hoc Committee on Measures of Postsecondary Preparedness will report its recommendations to the Governing Board no later than the November 2018 Board meeting.
To address its charge, the National Assessment Governing Board’s Ad Hoc Committee on Measures of Postsecondary Preparedness considered the trends that most likely will shape the future, and thereby shape the skills and knowledge that students will need to develop. Through meetings with expert panels and commissioning focused research papers, conducted with the support of its contractor HumRRO, the Committee pursued the answers to the following three research questions:

1. **Work of the future (readiness for what?):** What are we, as a nation, preparing students for? Changes in the workplace are not only inevitable, but are accelerating, driven by technological advances, demographic shifts, and social changes. The growing prevalence of self-driving vehicles, the more widespread use of robots, and advances in artificial intelligence are signs of existing innovations poised to dramatically change the jobs available to young Americans. Young Americans hold different expectations about work, and the ways in which people connect and communicate with each other are all changing. How will the workplace change given these emerging technologies? How will our communities change given these trends?

2. **Requisite skills for future work (skills for what?):** With a better understanding of the future workplace, we can better understand the skills that young Americans will need to succeed. But should we consider more than just workplace skills? What about skills like citizenship and financial literacy? How do these skills factor into the question of measuring postsecondary preparedness?

3. **Measures of preparedness (measures for what?):** Finally, what metrics exist to capture the skills that young Americans will need in the workplace, for their roles in their communities, and in their personal lives? Can such metrics include data from sources in addition to or instead of assessments? Additionally, what metrics do not exist but are needed to help the nation better understand if students are prepared as they exit high school, regardless of which paths they take—through college or other postsecondary learning experiences or directly to the workforce?

The following pages summarize the key findings from the various research and expert consultations; more detailed summaries are provided in the subsequent appendices.
Work of the Future – Readiness for What? (Q1)

Students beginning their K-12 education in 2017 are on course to graduate in 2030. Upon graduating high school, these students will face many choices for pursuing work and further learning. Traditionally, high school graduates have chosen between enrolling in postsecondary education or directly entering the workforce. Of 2017 high school graduates, for example, 67% enrolled in a postsecondary institution, 22% entered the work force, and 11% did neither.\(^1\) Postsecondary education has typically been a four-year college, two-year college, or technical school. Entry-level jobs are those that permit the employee to enter the workforce with little experience or education, often with the goal of moving onto a career path through valuable on-the-job experience. The military is an alternate pathway, providing training, experience, and a job; some choose to make serving in the military their career.

While similar options will remain for the graduates of the future, marked differences are anticipated as well. Current projections suggest that enrollment in degree-granting postsecondary institutions will increase by 2026\(^2\), whereas participation in the work force among 16-24-year old Americans will decline during a similar time frame\(^3\). The lines between postsecondary education and work may also become blurred, as universities and corporations expand their partnerships\(^4\) to offer new educational opportunities and employers begin to develop their own training and credentialing programs.\(^5\)

Students will also likely continue to move in and out of, and between, postsecondary institutions\(^6\) and work experiences; the curve and length of individual pathways will reflect personal needs, interests, and goals. And though they will vary in their timelines, each pathway will eventually lead to some form of work and for most a career, typically as part of the paid labor force. But what will the work of the future look like?

The graduates of 2030 will need to be prepared for a postsecondary experience that may look very different from that for which their parents and teachers were prepared. Although jobs that sound quite familiar to us now (e.g., teacher, veterinary technician, lawyer, engineer) will likely still be widely available, it is impossible to predict the range of specific jobs that will be available and sought-after more than a decade from now. Trends such as globalization, automation, and “big data” point to major changes in the world of work, both in terms of the available jobs and the work environment in which those jobs will be carried out.

---


Technology linking employers and employees will allow geographically dispersed people to vie for the same job, leading to an environment in which American workers will compete with others from around the globe.\footnote{Basu, K. (2016) \textit{Globalization of labor markets and the growth prospects of nations}. Retrieved from https://openknowledge.worldbank.org/bitstream/handle/10986/23929/Globalization00prospects0of0nation\_s.pdf?sequence=1&isAllowed=y} Once hired, they will be part of an increasingly dispersed and diverse work force. Advancements in communications technology will continue to bridge this distance, allowing for collaboration across space and time. Work of the future will often be contract-based, and workers for the future will need to be prepared to work as part of cross-cultural and mixed-discipline teams. With this increase in contract work, work environments will be more fluid, with individuals able to determine their own schedule and/or work space.

Automation will reduce the number of human workers needed to perform routine tasks in some fields, while creating new job descriptions in other fields and adding jobs directly related to the machines used to automate work\footnote{Atkinson, R. D., & Wu, J. (2017). \textit{False alarmism: Technological disruption and the U.S. labor market, 1850-2015}. Retrieved from: http://www2.itif.org/2017-false-alarmism-technological-disruption.pdf?\_ga=2.117549709.544738862.1522704813-61893732.1522704813}. Many customer service and middle management positions are expected to disappear as increasing numbers of transactions are completed via automated functions, thereby reducing the need for cashiers, clerks, and similar service providers, as well as those who would supervise them.\footnote{Andrew, P., Ip, J., & Worthington, J. (2014). \textit{Fast forward 2030: The future of work and the workplace}. Retrieved from file:///C:/Users/edickinson/Downloads/CBRE_Genesis_FAST_FORWARD_Workplace_2030_Full_Report_E.pdf} Jobs relating to transportation and logistics, office and administrative support, manufacturing, and service are also expected to decline due to increased automation.\footnote{Frey, C. B., & Osborne, M. A. (2013). \textit{The future of employment: How susceptible are jobs to computerisation?} Retrieved from: https://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf} At the same time, increasing numbers of workers will rely on artificial intelligence to assist them in their jobs. Humans and machines will collaborate to make decisions in the future world of work.\footnote{PwC. (2018). \textit{Workforce of the future: The competing forces shaping 2030}. https://www.pwc.com/gx/en/services/people-organisation/workforce-of-the-future/workforce-of-the-future-the-competing-forces-shaping-2030-pwc.pdf} Humans will also be responsible for designing and servicing these automated technologies.

Large quantities of data collected in real time will create job opportunities in data management and analytics, leading to expanded opportunities for those who can analyze and mine these data into information.\footnote{Ross, A. (2016). \textit{The industries of the future}. New York, NY: Simon & Schuster.} Data collection mechanisms will be seamlessly integrated into all parts of life, and jobs will require a broad range of employees to use data. Due to this unprecedented access to real-time data, organizations will change quickly to meet the demands of the markets in which they function.\footnote{Breu, K., Hemmingway, C., Bridger, D., & Strathern, M. (2002). Workforce agility: the new employee strategy for the knowledge economy. \textit{Journal of Information Technology} 17, 21–31.}

Job descriptions of the future will not be organized around clearly defined job titles, but rather around accumulated skills and experiences. Employers will leverage data to identify employees
who possess the experiences and skills specific to the job at hand.\textsuperscript{14} Potential employees will market themselves as uniquely qualified for the specifics of a project via online talent platforms and social media. Careers will be built out of freelance arrangements or other such “gigs”.\textsuperscript{15} Even those who opt for a more traditional career will likely hold around 12 jobs,\textsuperscript{16} either within one organization or across several organizations, over the course of their life. This will necessitate a continuous process of education and training throughout these future workers’ careers.\textsuperscript{17} Employers of the future will likely demand and incentivize on-going and just-in-time skill development to meet changing workforce needs.

Changes in the world of work have implications for postsecondary education as well. The graduates of 2030 will vary in terms of their high school experiences. Some will leave high school with college credits, Associate’s degrees and/or industry certifications, equipped with academic and job-specific knowledge and skills. Some will have paid work experience under their belts, in addition to or in lieu of service learning or other unpaid work experiences. Many will weigh decisions regarding what pathway is the best fit given their financial situation, their family’s needs, and their own goals, interests, and perceived abilities. Postsecondary institutions will need to use available data to make appropriate placement decisions, and then offer individualized instruction and a variety of student-focused services and support structures to ensure that students persist and ultimately graduate with competency in specific and generalized skills. Further, postsecondary institutions will need to adapt to meet the needs of an agile workforce that will be seeking opportunities for lifelong learning through additional credentials or courses designed to build new, or enhance existing, job skills.

**Requisite Skills for Future Work – Skills for What? (Q2)**

With this vision of the work of the future in mind, the skills needed to thrive in such an environment become apparent. Foundational academic skills, such as literacy and numeracy will continue to be valued and valuable in the postsecondary world of 2030. However, postsecondary success will also require a range of other cognitive, interpersonal and intrapersonal skills, as well as essential life skills. Cognitive skills include facility with technology, digital skills, computational thinking, and statistical literacy. Interpersonal skills include problem solving, communication, collaboration and cultural sensitivity. Intrapersonal skills include time management and flexibility. Essential life skills include financial and health literacy, as well as citizenship skills.

**Cognitive Skills**

Facility with technology will be in demand for jobs at all levels. Routine tasks formerly performed by humans may be taken over by robots or other technological advancements\textsuperscript{18}, leaving the human worker to service the technology or collaborate with the technology to complete more


\textsuperscript{17} PwC. (2018). Workforce of the future.

\textsuperscript{18} Ibid.
complex tasks. Employers will continue to increasingly seek individuals with computational thinking and digital skills to interact with data and new and emerging technologies.19

Basic digital skills will be essential for developing other skills. Education and on-the-job training will increasingly be delivered via digital platforms and will incorporate virtual reality simulations more frequently.20 More advanced digital skills include those required to create and use digital tools. The creation of digital tools, such as artificial intelligence and machine learning, requires STEM, analytic, and computational thinking skills. Employers will need programmers and innovators to develop new technologies to tackle more difficult challenges and improve collaboration, efficiency, and cost effectiveness.

Data management and analysis skills will also be in demand across job categories as data become more accessible and easier to distribute and share.21 Computational thinking and statistical literacy skills will be sought after as employers seek those who can effectively use, visualize and manipulate, and draw conclusions from data. Statistical reasoning skills will also be in demand as new technologies require humans who are able to train machine learning algorithms, explain how they work, and keep them operating.22

**Interpersonal Skills**

Regardless of the specific technical skill, workers of the future will not only need to possess such skills, but also will need to effectively apply those skills to real world problems. Problem-solving is an essential preparedness skill, one that will be required in more than one-third of jobs by the time the class of 2030 graduates high school.23 Employers and postsecondary educational institutions will seek graduates who are able to identify and select among many courses of action, and to do so in a dynamic work environment. Persistence in the face of uncertainty, and the ability to handle failure and identify next steps when things don’t go as planned, will also be essential for future postsecondary success.

Working as part of the dispersed and diverse workforce of the future will require communication and collaboration skills. Communication in the future world of work will require the ability to work with emerging communication technologies, along with the more traditional elements of communication such as listening and engaging in conversation. Active listening, in particular, is sought after by potential employers as it helps to create a positive work culture and supports collaboration, which in turn spurs innovation.24 Conversation skills are important because they

---

23 Thompson, C. (2016). *The top 10 skills that will be in demand by all employers by 2020.* Retrieved from https://www.businessinsider.com/wef-report-skills-workers-need-2016-1
contribute to an organization’s shared understandings, which may be critical for agile decision-making.  

Working collaboratively to solve problems will be an essential skill of the future, and will require building relationships, maintaining sensitivity to cultural differences, and seeing others’ perspectives. Forging positive relationships in the workplace will be key for an individual’s job satisfaction and will be an essential building block in the creation of a productive work environment. In the context of the workplace, cultural sensitivity includes working effectively alongside someone from a different cultural background who may approach workplace behaviors differently. Coworkers from different cultural backgrounds may engage in different behaviors and hold different work-related values, and culture-based misinterpretations can have implications for the success of collaborative efforts. Perspective taking involves awareness of others, regulating one’s emotions and empathy, and correctly interpreting what others are trying to communicate. 

**Intrapersonal Skills**

The fluidity of work arrangements in the future will necessitate time management skills, as individuals work on multiple project teams or juggle multiple gigs. Time management requires a variety of skills: estimation of effort, scheduling, prioritizing, delegation, and monitoring a to-do list, among myriad others. An individual with strong time management skills can not only project the amount of time and effort a given task will require, but also inhabit the mindset to meet deadlines and, perhaps as importantly, recognize when a deadline cannot be met and adapt accordingly. While employers have historically valued employees with solid time management skills, in the expanding gig economy the individual entrepreneur’s personal success will depend upon it. 

Flexibility and adaptability will also be essential skills, whether as an independent contractor providing services to multiple clients or as a career employee adjusting to changing roles and expectations in a dynamic environment. In fact, flexibility and adaptability, particularly in the context of interpersonal communication, are among the uniquely human skills that may prevent some occupations from becoming fully automated. In addition to being key to specific careers, flexibility will also be integral to managing the projected evolving career path an individual will undertake over the course of a working lifetime. As companies demand upskilling or reskilling, the adaptable employee will be at an advantage. 

**Life Skills**

Finally, graduating high school students will be best prepared for postsecondary success if they have acquired essential life skills. Skills such as financial literacy, and health and wellness literacy will help to ensure that decisions are made with long term implications in mind, setting the stage for continued success. Finally, high school graduates of 2030 will require citizenship skills, understanding the relations between individuals and society, the organization of the state communication.

---


and how democracy functions, and the roles and responsibilities of the individual in the world around them.\textsuperscript{28}

**Measures of Preparedness – Measures for what? (Q3)**
Postsecondary preparedness is a culmination of in-school and out-of-school experiences leading up to high school graduation. It is a multi-faceted concept that will require a multidimensional measurement approach. Currently, there is no uniform and comprehensive measure of postsecondary preparedness. NAEP measures academic preparedness for college without remediation with its reading and math assessments, but this does not address the wide range of skills discussed above. Existing measures such as NAEP assessments of academic skills will need to be combined with other existing data sources and will likely need to draw on new data sources or innovative measurement approaches, to gain a full picture of the state of postsecondary preparedness.

NAEP assesses a wide range of content areas at grade 12, including civics, economics, foreign language, geography, math, reading, science, technology and engineering literacy, U.S. history, and writing. The NAEP civics frameworks, for example, outline “intellectual skills” such as analyzing and evaluating, taking, and defending positions, along with participatory skills such as interacting, monitoring, and influencing.\textsuperscript{29} These intellectual skills may align with several interpersonal skills outlined above (e.g., communication, collaboration). Similarly, the technology and engineering literacy assessment frameworks outline three practices (understanding technological principles, developing solutions and achieving goals, and communicating and collaborating) that may reflect some of the described technical, intrapersonal, and interpersonal skills.\textsuperscript{30} NAEP also administers complex and interactive item types that may tap into skills beyond the academic skills measured. For example, a reading item may tap into cultural awareness skills, or a writing item may tap into effective communication skills. Finally, students who participate in NAEP are also encouraged to complete a student survey, which collects information about students' educational experiences.\textsuperscript{31}

NAEP also collects data as part of the High School Transcript Study. This study is designed to reflect a nationally representative sample of U.S. schools and a sample of students that is representative of the graduates from each participating school.\textsuperscript{32} Data collected include the courses taken during high school, high school credits earned (including those earned in middle school), and final course grades. Information about course-taking patterns are not only informative about academic skills to which students have been exposed, but also provide indirect information about other important skills such as persistence and time-management.

Other data sources outside of NAEP, such as data regularly collected by or stored at state education agencies, may also be useful in informing our understanding of postsecondary preparedness. Participation in education and work experiences, from course attendance, to maintaining employment, to engaging in volunteerism or service learning experiences can

provide evidence of individuals’ development of important preparedness skills that can be aggregated to higher levels. Data embedded in these education and work experiences are another potential source of preparedness data. For example, experiences with technology, individual portfolios, or team projects are elements of applied educational and work experiences that yield information about relevant skills.

Data from assessments administered at the state level could also be mined for relevant information. For example, performance tasks designed to reflect real-world contexts may provide insight into blends of skills, such as problem-solving within an academic content area. Computer-based assessments could be a source of data on digital skills. Performance-based assessments may tap into persistence, problem solving skills, and analytical thinking skills.

New and emerging technological resources present additional opportunities to gather data from education and work experiences. Micro-credentials, also known as digital badges, are earned upon completion of a short course, administered online or in a more traditional classroom setting. In the future, blockchain technology may allow parties to record transactions and maintain a permanent digital record of them, which could be used to document achievements within a particular organization or program.33

**Current Context/National Need**

NAEP has been a leader in understanding and communicating what the nation’s students know and can do for decades. It has recently been argued that the timing is right for expansion of 12th grade NAEP testing, citing the need for trustworthy data on college and career readiness, and for a data source that would allow states to compare themselves to one another and to the country as a whole.34 With its well-established infrastructure, NAEP is in a unique position to gather postsecondary preparedness data at the state level. NAEP data, along with other NCES data sources, will inform those interested in postsecondary preparedness policy, and will be useful to states as they work with industry and business partners to expand job opportunities, and to support local-level efforts to prepare students for life after high school. These data will also be of use to postsecondary education providers and employers as they seek to ameliorate gaps in preparedness, and to inform the public about how prepared high school graduates are for their next steps.

Over half of states have developed definitions of *college and career readiness*.35 In a large majority of these states, a single definition is used to describe readiness for college and readiness for career, and some address preparedness for postsecondary life more broadly. Some of these definitions focus on academic skills necessary to enter credit-bearing college courses without remediation or to successfully complete postsecondary job training program that will lead to a career that can support a family. Others describe additional skills beyond the academic, such as communication and collaboration. A very small number also include activities such as lifelong learning and civic engagement. Although there has been a substantial amount of work done at the state level, there are still states that have not developed a formal definition. Further, the existing state-level definitions are varied enough to make comparisons among the states confusing. NAEP is in a unique position to lead the charge in collecting data to inform a

---

common understanding and, by extension, identifying the most essential elements of postsecondary preparedness.
Notes of the Expert Panel Meeting Representing Industry
February 22, 2018

National Assessment Governing Board
Ad Hoc Committee on Measures of Postsecondary Preparedness

As part of meeting the charge of the Ad Hoc Committee on Measures of Postsecondary Preparedness, HumRRO organized and facilitated a meeting with industry experts. The purpose of this meeting was to get input from leaders and experts in industry about (a) the jobs that will exist in 2030, (b) the skills that these jobs will require, and (c) the measures/indicators that would be needed to provide a status of elementary and secondary students with respect to these skills.

We were fortunate to assemble an exceptional panel of experts and leaders. The panel members included Ms. Paula Collins, Texas Instruments, Mr. Marcelino Ford-Livene, Intel Corporation, Dr. Scott Heimlich, Amgen Foundation, Dr. Chauncy Lennon, JPMorgan Chase, and Mr. Reginald McGregor, Rolls-Royce Corporation.

The meeting was held on February 22, 2018 in Alexandria, Virginia. An overview of the National Assessment Governing Board and the charge of the Ad Hoc Committee on Measures of Postsecondary Preparedness, along with the agenda and logistical information for the meeting were sent to the panelists in advance.

Thanos Patelis (HumRRO) opened the meeting and after quickly informing the group of some logistics, Terry Mazany provided an overview and led the attendees through introductions. Then, Thanos Patelis facilitated the meeting around the three areas of inquiry involving (a) the jobs of 2030, (b) the skills that they will require, and (c) the measures/indicators that will be important to provide. Finally, Terry Mazany offered some concluding comments. The agenda and the list of all attendees is in Appendix A.

The purpose of this document is to provide information on the themes and comments made by the panelists. The information in this report is meant to provide insight into the rich conversation and comments provided by the expert panelists.

The Future of the Workplace and Work

- The titles of the jobs in 2030 cannot be predicted. However, the jobs of the future will require many skills and will be driven by globalization, artificial intelligence, and “big data”.
  - Globalization will change the workplace, from the types of jobs available (i.e., global competition for jobs) to working on cross-cultural teams.
  - Workplace integration will increase (e.g., working across disciplines instead of in silos by discipline).
  - The pace of automation and existence of the internet enable rapid access to information which will affect what employees do on the job and their job descriptions. The use of the internet and automation will only increase.
  - Employers should embrace new methods of communication, driven by the next generation. For example, hiring managers may not be familiar or may be uncomfortable with the latest communication modes of those applying for jobs. Rather than allowing that to impact negatively on job applicants, employers should
acknowledge the differences as innovation or trends to monitor. Job applicants may also need to be attuned to this dynamic.

- Technology will be at the forefront. For example, JP Morgan Chase is a “tech company that also loans money”; they do not consider themselves primarily a financial institution.
- Complicated tasks can be handled by automation (which will replace some jobs). Employees of the future will need to work with automated equipment and employees will be needed to design and service the automation.
- Complex tasks will take human thought (and these types of jobs will remain and additional ones will be added in the future).

- There is and likely there will continue to be a duality in the job descriptions of the future: academic skills and college degree required versus high school diploma and training and apprenticeship experience required. Panelists noted they come from the academic skills track and although they acknowledge the diploma-training track, they suggested consulting with experts in that area for a more detailed picture of what the future holds for those not following the 4-year college track.
  - Need to hire the person with the right skill set, not the person with the most qualifications (who may be overqualified and a poor fit for the work). This is sometimes a tendency when college-graduate hiring managers put more emphasis on college degree, the background they come from and perspective they bring to their job, than is warranted by the demands of the job being filled.
  - Most jobs that do not require a 4-year college degree, will require additional training, such as a 2-year college degree, technical training, or post-secondary education and/or training leading to certification.
  - Employer provides job skills (e.g., specific knowledge and procedures), while employee brings workplace competencies to the job (see competencies in the skills needed in the future). More job-related training will be provided by the employer, such as in-house mini-MBA programs provided by large corporations.
  - Continuous learning will be required to keep up with change. The employer will support or provide the training or education; the employee must participate to keep pace.

- Panelists indicated the need for initiatives to empower students, especially those who are “at-risk” and do not have role models, with an understanding of the labor market and expose them to employment options. Suggestions for empowering students so they are ready for post-secondary steps to meet their goals:
  - Help them define pathways to jobs.
  - Assist in setting goals; define an individual’s “north star”.

- Employer/employee relationships will change.
  - More contract work will emerge, which allows workers to dictate own schedule and/or workplace.

- Office space will be different.
  - For example, if employees come to the office, they will use a laptop and choose a work space area plugging into the network. The exact location may vary and will be more fluid than today.
Skills Needed in the Future

- Panelists described the need for employees to be able to apply skills, which defines competencies. Having a skill is not sufficient. Must know how to apply the skill to real world problems.

- The skills that were highlighted were as follows:
  - Ability to collaborate with people and machines, as the workplace incorporates more technology and automation as well as more collaboration.
  - Ability to interact with technology in jobs at all levels. Career Technical Education (CTE) can provide skills and certification for certain jobs.
  - Data skills are in demand - data is the new oil.
  - Less focus on job-specific content skills and more on workplace competencies:
    - Critical thinking, effective communication, collaboration, adaptability, problem solving, creativity, integrity, community/workplace citizenship, agility, learning disposition, persistence, attitude, interest.
  - Able to handle failure – know what to do when the button fails.

- Need power skills and experience, especially for at-risk students, to navigate the job market and succeed in entry-level positions – resume writing, oral communication, working on teams, basic reading/writing and mathematics ability.

Measures of Skills in the Future

- Consider measuring post-secondary readiness skills in grade 8.
- Maintain traditional knowledge measures (i.e., reading, mathematics).
  - Some went as far as to say that these measures of academic skills should not be removed and any other measures should be added.
- Design-build skills can be measured by persistence. Do you persist until object is built?
- Measure application of skills at grade 12. Can students demonstrate their skills (versus showing their knowledge of skills)?
- Add new measures tapping workplace requirements. Be creative in measuring skills (e.g., use certificates or credentials). Leverage CTE curriculum and measures.
  - In the interview process for candidates, hiring managers will give a problem to solve. Therefore, such metrics that demonstrate process and results of solving problems would be helpful.
- Need measures on collaboration, empowerment, and creativity.
- Tie relevancy of measures to industry and align with education. Do this regionally so that measures of preparedness are informative to:
  - students (do they have the skills needed for jobs in their community?),
  - industry (do local job applicants have the skills needed for jobs being offered in their community?),
  - educators (are they preparing students for post-secondary opportunities in their community?), and
  - policy makers (does the local workforce have the skills that industry in their community require?).
- While this may not be the Governing Board’s responsibility, students should be given the ability to develop digital portfolios, including coursework and experiential activities, in school to demonstrate their skills and achievements. This would be helpful to employers.
- The measures must keep evolving as the type of work and required skills change over time.
One interesting observation was that the panelists described job training interventions for at-risk youth with measures of program success embedded as artifacts of the experience. Did the participant build something? While the final product might not have been their initial design, the focus was on the creative process and the ability to troubleshoot problems as well as to persist in developing the final product.
Appendix A: Meeting Agenda and Attendees

Expert Panel Meeting
National Assessment Governing Board
Ad Hoc Committee on Measures of Postsecondary Preparedness

February 22, 2018 | Agenda

11:00 to 11:05 AM  Start Meeting
Thanos Patelis, Facilitator, HumRRO

11:05 to 11:15 AM  Welcome and Introductions
Terry Mazany, National Assessment Governing Board Member
Chair, Ad Hoc Committee on Measures of Postsecondary Preparedness

11:15 AM to 12:00 PM  Work of the Future
Thanos Patelis, Facilitator, HumRRO

Guiding Questions:
- What do you see as the type of jobs graduating high school seniors will have in 2030?
- Compared to jobs now, what kind of trends do you see emerging for jobs in 2030?
- Do you foresee any differences of jobs by industry or do you expect similar trends to occur for all jobs?
- What do you see as expectations of employers for these students?
- How do you envision the hiring process to be?
- What role will postsecondary institutions play in training and preparing students for these jobs?

12:00 to 12:15 PM  Break to get lunch

12:15 to 1:00 PM  Skills for the Work of the Future
Thanos Patelis, Facilitator, HumRRO

Guiding Questions:
- What types of skills will graduating high school seniors need to have in 2030 in order to get the jobs in 2030?
- What would you consider pre-requisite skills vs. skills that can be acquired on the job?
- What role will postsecondary institutions play in training these skills?
- What would a hiring manager in 2030 look for in prospective hires?

1:00 to 1:45 PM  Measures of these Skills Associated with Work of the Future
Thanos Patelis, Facilitator, HumRRO

Guiding Questions:
- What measures do you see being used to represent these skills?
- What metrics would provide helpful information in the aggregate about the skills of graduating high school seniors?

1:45 to 2:00 PM  Final thoughts and concluding remarks
Terry Mazany, National Assessment Governing Board Member
Chair, Ad Hoc Committee on Measures of Postsecondary Preparedness
Attendees

Expert Panelists:
- Paula Collins, Texas Instruments
- Marcelino Ford-Livene, Intel Corporation
- Scott Heimlich, Amgen Foundation
- Chauncy Lennon, JPMorgan Chase
- Reginald McGregor, Rolls-Royce Corporation

Governing Board Members:
- Terry Mazany, Chair, Ad Hoc Committee on Measures of Postsecondary Preparedness
- Honorable James E. Geringer, Former Governor of Wyoming, Cheyenne, Wyoming
- Carol Jago, Associate Director, California Reading & Literature Project at UCLA, Oak Park, Illinois
- Dale Nowlin, Teacher and Mathematics Department Chair, Bartholomew Consolidated School Corporation, Columbus, Indiana
- Honorable Beverly Perdue, Former Governor of North Carolina, New Bern, North Carolina
- Linda P. Rosen, Chief Executive Officer, Change the Equation, Washington, DC
- Chasidy White, Director of Strategic Initiatives, Office of the Superintendent, Montgomery, Alabama

Governing Board Staff Members:
- Bill Bushaw, Executive Director
- Lisa Stooksberry, Deputy Executive Director
- Lily Clark, Assistant Director for Policy & Research
- Laura LoGerfo, Assistant Director for Reporting & Analysis
- Munira Mwalimu, Executive Officer & Contracting Officer
- Sharyn Rosenberg, Assistant Director for Psychometrics
- Angela Scott, Management & Program Analyst

HumRRO Staff Members:
- Monica Gribben, Senior Staff Scientist
- Deirdre Knapp, Vice President, Assessment and Evaluation in Education and the Workplace
- Jackson Millard, Research Associate
- Thanos Patelis, Principal Scientist
Expert Panelists

Paula Collins  
Vice President, Worldwide Government Relations  
Texas Instruments

Paula J. Collins is vice president of Worldwide Government Relations for Texas Instruments where she leads the Company’s advocacy activities in the United States and abroad. She joined Texas Instruments in 1999 as Director of Government Relations and managed the Company’s legislative and public policy activities on a wide range of issues, including immigration, funding for basic research and education.

Ms. Collins came to Texas Instruments with extensive government, corporate and business association experience. After serving as a legislative assistant on Capitol Hill, she joined American Express Company, where for ten years she directed the Company’s legislative activities on a wide range of public policy issues including a number of trade initiatives. In 1993, she joined the Business Roundtable where she worked closely with corporate leaders to develop and implement public policy campaigns on international trade, budget and workforce initiatives. From 1995-1997, she directed international trade relations at Eastman Kodak Company and from 1997-1999 was a principal with The Fratelli Group, a strategic communications firm where she played an active role in the development and implementation of comprehensive public affairs strategies for several coalitions on trade and telecommunications issues.

Ms. Collins is a graduate of Yale University and attended the Program for Management Development at Harvard Business School. She is an active participant in her church and local civic organizations, and is a member of several professional organizations. She is a member of the Board of Directors and Executive Committee of the Information Technology Industry Council, and chairman of the Board of the Task Force on American Innovation.
Marcelino Ford-Livene  
General Manager, Global Programs and Alliances  
Intel Corporation

Marcelino Ford-Livene is the General Manager of Global Programs and Alliances for Intel’s Worldwide Corporate Affairs Group. In this capacity, he leads the organization charged with designing the framework and strategic plan for identifying and prioritizing win-win strategic alliances, relationships and partnerships with various global industry, government and special interest groups that advance the strategic direction of Intel’s Diversity and Inclusion Initiative. Prior to this role, Ford-Livene was the General Manger of New Channels and Advanced Advertising for Intel Media, where he led the organization charged with programming, licensing and distributing new format television channels and advertising-supported video-on-demand programming. He was also responsible for advertising sales, advertising operations, audience research and data analytics for Intel Media’s OTT services. He also co-authored patents on TV viewership analytics and advanced advertising behavioral targeting. Prior to Intel, he was a senior member of TV Guide’s corporate development and planning team. He has also held senior positions with the U.S. Federal Communications Commission in Washington, DC. He served as Special Counsel for New Media Policy for Chairman William E. Kennard and as Senior Counsel and Director of Media Strategic Analysis for the FCC’s Office of Strategic Planning under Chairman Michael Powell. Ford-Livene was the Division Chairman of the Interactive Media Division for the American Bar Association’s Forum on the Entertainment and Sports Industries from 2006 to 2013. He also served for eight years on the board of the TV Academy, the organization that awards the prestigious Primetime Emmy for creative excellence in the television industry. He was also the TV Academy’s Board Secretary and a member of its Executive Committee from 2010 to 2013. He is currently the Co-Chairman of the TV Academy’s Diversity Committee and a founding board member of the Digital Diversity Network. Corporate boards that Ford-Livene has served on include Delivery Agent in San Francisco, CA and TRA Global, which was acquired by TiVo. Ford-Livene earned a B.A. in economics from UC San Diego, a J.D./M.B.A. from the University of Illinois and has completed an Executive Leadership Program at Harvard Business School.
Scott Heimlich  
Vice President, Amgen Foundation

Scott M. Heimlich is vice president of the Amgen Foundation. He is responsible for the strategic management and direction of the Foundation’s science education portfolio, including the development and oversight of key initiatives at the K-12 and higher education levels. He was the principal architect and continues to lead the Amgen Scholars Program, the Foundation’s largest initiative providing undergraduates with access to research opportunities at premier educational and research institutions across the world. Under his leadership, the Amgen Biotech Experience transformed from a local program into a multi-site, international initiative bringing biotechnology lab experiences to over 80,000 secondary students a year. With these and many other initiatives, the Foundation’s commitment to science education recently surpassed the $125 million milestone.

Prior to joining Amgen in 2005, he served in positions at the University of California, Los Angeles, Los Angeles Pierce College, University of Southern California, and a junior high school in Japan. He holds a bachelor’s degree, master’s degree, and doctorate in education from the University of California, Los Angeles.

Chauncy Lennon  
Managing Director and Head of Workforce Initiatives  
Global Philanthropy  
JPMorgan Chase & Co.

Chauncy Lennon leads JPMorgan Chase & Co.’s initiatives to promote economic opportunity through investments in workforce practice, innovation, and policy. These include New Skills at Work, a $250 million global initiative to support demand-driven workforce systems that promote prosperity for workers and industries; New Skills for Youth, a $75 million initiative to increase the number of young people who complete career pathways that begin in high school and end with postsecondary degrees or credentials aligned with good-paying, high-demand jobs; The Fellowship Initiative, a program providing young men of color with learning experiences that help them achieve their education and career potential; and a $17 million investment in Summer Youth Employment Programs in US cities to help underserved youth obtain the skills necessary to build lasting careers.

He serves on the New York City Workforce Development Board, the College Promise Campaign Advisory Board, and the Neighborhood Trust Financial Partners Board.
He joined JPMorgan Chase from the Ford Foundation, where his grant-making focused on promoting economic advancement for low-income workers by improving access to workforce development and work support programs. Prior to the Ford Foundation, he was senior vice president for Asset Building at Seedco, a national workforce development intermediary. He also has extensive experience researching the mobility patterns of the working poor. He earned his Ph.D. in anthropology from Columbia University, master's degree from the University of Chicago and bachelor's degree from Williams College. He has taught urban studies at Columbia's School of International and Public Affairs and Barnard College.

Reginald McGregor
Manager, Research & Technology Strategy Group
Rolls-Royce Corporation

Reginald McGregor, Manager of Engineering Employee Development and STEM Outreach at Rolls-Royce Corporation. He is a Mechanical Engineer with over 15 years' experience in various engineering roles. He spent over 8 years in early career development managing the engineering co-op; high school internship and graduate development programs. Reginald holds BS in Mechanical Engineering, MBA and currently completing a MS in Technology Leadership and Innovation. He is very active in workforce development and STEM education and serving the community. Reginald enjoys reading, outdoor activities and spending time with family.

Reginald serves on several boards and committees including the Governor-appointed Region 5 Works Council, President of the Lawrence Township School Board, Indiana STEM Advisory Council, STEMx National Advisory Board, Purdue Engineering Education Industrial Advisory Council, Marion County Superintendents STEM Coalition, Indiana Chamber of Commerce K-12 and Workforce Committees, Million Women Mentor Steering Committee, Indiana Afterschool Network Board, and EmployIndy Youth Committee.
As one step in addressing the charge of the Ad Hoc Committee on Measures of Postsecondary Preparedness, HumRRO organized and facilitated a meeting with a select group of higher education innovators. The purpose of this meeting was to elicit input from leaders and experts in higher education about (a) the jobs that will exist in 2030, (b) the skills that these jobs will require, and (c) the measures/indicators that would be needed to determine the status of elementary and secondary students with respect to these skills.

We were fortunate to assemble an exceptional panel of experts and leaders. The panel members included Dr. Sarah DeMark, Vice President of Academic Programs, Western Governors University; Dr. Pradeep Kotamraju, Bureau Chief, Career and Technical Education, Division of Community Colleges and Workforce Preparation, Iowa Department of Education; Mr. Michael Morsches, Dean of Learning Enrichment and College Readiness, Moraine Valley Community College; Dr. Yvette Mozie-Ross, Vice Provost for Enrollment Management and Planning, University of Maryland, Baltimore County; and Dr. Holly Zanville, Senior Advisor for Credentialing and Workforce Development, Lumina Foundation. Also, in attendance were some Governing Board members, Governing Board staff members, and HumRRO staff, listed in Appendix A.

The meeting was held on April 19, 2018 in Chicago, Illinois. An overview of the National Assessment Governing Board and the charge of the Ad Hoc Committee on Measures of Postsecondary Preparedness, along with the agenda and logistical information for the meeting were sent to the panelists in advance of the meeting.

Thanos Patelis (HumRRO) opened the meeting and after quickly informing the group of some logistics, Terry Mazany, Ad Hoc Committee Chair, set the stage for the role of NAEP in the future, given the impact of technology on work as well as the economic and global context in which students enter the post-secondary world. He led the attendees through introductions. Thanos Patelis facilitated the meeting around the three areas of inquiry involving (a) the jobs of 2030, (b) the skills these jobs will require, and (c) the measures/indicators needed to measure these skills. Finally, Terry Mazany offered some concluding comments. The agenda and the list of all attendees is in Appendix A.

The purpose of this document is to summarize the themes and comments made by the panelists. The information in this report is meant to provide insight into the rich conversation and comments provided by the expert panelists.
The Future of the Workplace and Work

With experts representing higher education, the discussion of the future of the workplace and work focused on pathways to work, primarily through postsecondary education and training.

- Postsecondary institutions need to create pathways to develop agile employees who are open to lifelong learning.
- Lifetime or continuous learning will become the norm. Employees will need to continue to learn from different providers, from colleges/universities to specific training courses to experiential opportunities, throughout their lives. Information technology (IT) workers already face this with a variety of certifications for specific technology tools and applications. Highly-regulated occupations will likely be the last ones to make changes.
- Postsecondary institutions need to partner with employers to identify education and training needs so that graduates possess the knowledge and skills needed for jobs.
  - Look to IT which is leading the way in defining job requirements and credentials for employees.
  - One of the panelists described a keynote presentation by the CEO from Chegg, Dan Rosensweig, describing the current disconnect between expectations and responsibilities of employers, higher education, and students. He illustrated this by placing each of the stakeholders at the vertices of a triangle with arrows facing outward indicating a lack of working together rather than arrows pointing inward, toward each other, signaling collaborative planning and working together toward similar goals.
  - Educators can be resistant to business models.
- There are still barriers to postsecondary education. Although community colleges have an open policy (in some states students do not need a high school diploma to enroll in community college), students may find it difficult to pursue their desired major or to matriculate. Prerequisites and competitive admission in selected programs (e.g., healthcare) are barriers to entry.
  - Similarly, some 4-year colleges guarantee admission to those with associate’s degrees, but cannot guarantee admission into specific programs due to enrollment capacity and accreditation requirements such as completion of specific coursework.
  - Some community college graduates are not prepared for 4-year colleges and universities because their 2-year institutions have limited qualification requirements for instructors and low standards for their graduates. Both of these factors could be a barrier to continued education.
- More individualization in postsecondary education requires “policy by anomaly.”
  - In developmental education, need to identify what students need and how to get it to them. Placing students on paths matching their goals raises retention rates.
- Strong partnerships are needed between 2- and 4-year institutions of higher education to facilitate students’ transfer between schools.
  - High school graduation projections show Hispanics are the fastest growing group and many of this group begin their postsecondary studies in community college.
  - Many students are graduating from high school with associate’s degrees obtained through early middle college programs and dual enrollment.
- Colleges and universities must provide different, perhaps individualized, services to students who enter at different points on the pathway to a 4-year degree. Historically, 18-year-old high school graduates enter as freshmen with new-student services and support structure

---

for the first year or two. Institutions are now called on to help a select group of high school graduates entering college with associate’s degrees, yet perhaps still needing wraparound services due to their youth (compared to the services offered to 20-year-old or older students transferring to a 4-year program with an associate’s degree). Other students may start and stop their education multiple times and attend several institutions before graduating.

- To prepare students for future jobs, we need vertical and horizontal articulation. For horizontal articulation, students need technical, academic, and employability skills (e.g., grit, self-understanding). For vertical articulation, the key is determining at what age/grade to start. High school staff say it needs to start in middle school; middle school staff say it needs to start in elementary school.
- Need a mechanism to validate training and experience as part of the pathway to a degree. More and more high school graduates are already working through the gig economy. Other students have jobs and families while attending college.
  - Look to the military; they validate training as credits.
  - Western Governors University (WGU) provides micro-credentials or badges as students achieve milestones to show them the skills and knowledge attained as they work toward their bachelor's degree.
  - Give students the ability to curate their work and educational experiences.
- There is tension between an integrated approach providing a broad range of skills (academic, technical, and employment-oriented) and the business need for a narrow, specific set of skills to meet a skill shortage. One is too esoteric, the other too pragmatic.
- Post-secondary institutions will not be the destination, but a vehicle for certifying student competencies.
- Expect the acquisition and use for knowledge and skills to flip. Currently, knowledge is the base foundation provided by formal education and we obtain skills as needed. In the future, skills will be the base and we will obtain knowledge as needed.

Skills Needed in the Future

- Don’t teach students to do what a robot can do better.
  - Robots are better than humans at pattern recognition, repetitive tasks, etc. but they are not able to understand nuance of language, social relationships, or creativity.
  - It will be important for humans to connect domains.
  - McKinsey has developed a list of human skills such as empathy, planning, creativity, common sense, sense making, novel thinking, nuance of language, social relationships, etc.37
- In addition to content or professional knowledge, students need:
  - practical transition skills
  - key learning skills and cognitive strategies
  - strong foundation of self-understanding and engagement strategies
  - critical thinking
  - affective mindset and skills
  - meta learning
  - financial literacy
  - information technology literacy
  - health and wellness literacy.
- Schools can provide learning and workplace skills.

---

College experience courses for high school students.

WGU offers eight synchronous online sessions with a small, facilitated cohort on skills such as self-efficacy, communication, and learning styles. In a pilot test with at-risk students, there were significant positive outcomes: performance in courses as well as retention increased. Some of the skills, including leadership and communication, were identified by the medical profession as ones missing in graduates. These skills not only make graduates better job candidates but also more resilient students.

- Consider where or why skills are needed to build awareness of how skills fit into work.
- Four-year institutions look for grit or persistence as a necessary skill for student success. Students with a solid academic foundation and grit should be able to succeed, whereas students with a strong foundation of academic knowledge and no grit may not be able to handle the rigor of college.
- Class attendance is the best predictor of success, as evidenced both by anecdote and research. Some colleges require attendance and initiate interventions if students do not attend class.
  - There is a question of how to measure attendance for online courses. One approach is to look at student engagement using interaction data from Learning Management Systems (LMS).
- Employers need to learn how to get “unstuck” when in a challenging situation.
- Employers are looking for people who can work across left and right brains and are able to work with technology.

**Measures of Skills in the Future**

- Employers offer performance-based pay for high-value, high-priority credentials supporting ability to use skills.
  - Students may demonstrate their skills through portfolios.
  - Use blockchain[^38] to document achievements and portfolio.
- Need new types of student assessment.
  - Current assessments focus too much on knowledge and not enough on skills, character, and meta learning.
  - Students take most current assessments working alone rather than in teams. Need authentic assessments of teamwork with hands-on performance components.
- Leading-edge assessments use simulation and are more applied, with problem solving scenarios that assess whether you can use knowledge.
- Create dashboards for parents and students to see skill attainment, including credentials.
- Use micro credentials and then stack those credentials to meet employer-relevant needs.
- There is a tension between broad versus specific measurement of skills.
- Include all stakeholders in identifying what and how to measure skills.
- Measuring college or postsecondary readiness is different than college or postsecondary success.
- Some postsecondary institutions use transcripts, others don’t.
  - Transcripts could provide an opportunity to leverage high school data for postsecondary instructors to know what students have done prior to college and to personalize postsecondary instruction.
  - Expect seat time to be a less helpful measure from an industry perspective. They will be interested in a “transcript” with learning opportunities, perhaps using blockchain technology.

[^38]: For information about blockchain: [https://hbr.org/2017/01/the-truth-about-blockchain](https://hbr.org/2017/01/the-truth-about-blockchain)
For transcripts to be useful to instructors, need a way to standardize them.
Need to include attendance on transcript.
- Metrics of academic rigor exist with validity evidence provided to support their value in predicting college outcomes.
- Concern with the shelf life of measures such as SAT or ACT, course grades, etc. Are high school results as valid for older, returning students?
- Metrics should include student employment.
- Measures of service learning are needed.

**Reflections**

Terry Mazany offered four reflections on the discussion:
1. We need to project all of the allied trends in society to 2030. Work is shifting to a gig economy. This will be the reality for 16- to 18-year-olds in 2030. We need to factor the expected changes in the economy of 2030 into the skills required to work in the future. Data is the new oil. Micro-credentialing and digital badges will more and more populate transcripts and portfolios.
2. There will be several paradigm shifts: (a) knowledge/skill flip, (b) everything has a developmental progression except technology, (c) the nontraditional student of today will be the traditional student of tomorrow, (d) students will be agents for themselves, and (e) a world where trust is collapsing in every venture except nonprofit ventures – blockchain as a key to build this trust.
3. We are in-between systems. We need to maintain an ecological perspective of each part of the system and look at the reciprocal changing role of employers.
4. The role of NAEP: We need to align NAEP with the requirements of Every Student Succeeds Act (ESSA), such as conditions of learning. This might be done by back-mapping the requirements of ESSA with what NAEP provides.
Appendix A: Meeting Agenda and Attendees
Expert Panel Meeting
National Assessment Governing Board
Ad Hoc Committee on Measures of Postsecondary Preparedness

April 19, 2018

11:00 to 11:05 AM  Start Meeting
Thanos Patelis, Facilitator, HumRRO

11:05 to 11:15 AM  Welcome and Introductions
Terry Mazany, National Assessment Governing Board Member
Chair, Ad Hoc Committee on Measures of Postsecondary Preparedness

11:15 AM to 12:00 PM  Work of the Future
Thanos Patelis

Guiding Questions:
- What do you see as the postsecondary pathways that high school seniors graduating in 2030 will be choosing among? (11:15-11:40)
- Compared to now, what kind of trends do you see shaping postsecondary education in 2030? (11:40-12:00)

12:00 to 12:15 PM  Break to get lunch

12:15 to 1:00 PM  Skills for the Work of the Future
Thanos Patelis

Guiding Questions:
- How have postsecondary entrance expectations changed in recent years? (12:15-12:40)
- What types of competencies and content knowledge will graduating high school seniors need to be prepared for postsecondary pathways in 2030? (12:40-1:00)

1:00 to 1:45 PM  Measures of these Skills
Thanos Patelis

Guiding Questions:
- What measures do you see being used for these competencies? What will require new or updated measurement tools? (1:00-1:20)
- What metrics would provide helpful information in the aggregate about the competencies of graduating high school seniors? (1:20-1:45)

1:45 to 2:00 PM  Final thoughts and concluding remarks
Terry Mazany
Attendees

Expert Panelists:
- Sarah DeMark, Vice President of Academic Programs, Western Governors University
- Pradeep Kotamraju, Bureau Chief, Career and Technical Education, Iowa Department of Education
- Michael Morsches, Dean of Learning Enrichment and College Readiness, Moraine Valley Community College
- Yvette Mozie-Ross, Vice Provost for Enrollment Management and Planning, University of Maryland, Baltimore County
- Holly Zanville, Senior Advisor for Credentialing and Workforce Development, Lumina Foundation

Governing Board Members:
- Terry Mazany, Chair, Ad Hoc Committee on Measures of Postsecondary Preparedness
- Dale Nowlin, Teacher and Mathematics Department Chair, Bartholomew Consolidated School Corporation, Columbus, Indiana
- Alice Peisch, Legislator, Massachusetts House of Representatives, Wellesley, Massachusetts
- Chasidy White, Director of Strategic Initiatives, Office of the Superintendent, Montgomery, Alabama

Governing Board Staff Members:
- Bill Bushaw, Executive Director
- Lisa Stooksberry, Deputy Executive Director
- Lily Clark, Assistant Director for Policy & Research

HumRRO Staff Members:
- Monica Gribben, Senior Staff Scientist
- Sunny Becker, Principal Staff Scientist
- Thanos Patelis, Principal Scientist
Sarah DeMark, Ph.D.
Vice President of Academic Programs
Western Governors University

Sarah DeMark joined nonprofit Western Governors University (WGU) in September 2014, and serves as the Vice President of Academic Programs, responsible for leading WGU’s portfolio strategy as well as the design and development of the university’s competency-based degrees, curriculum and assessments. This portfolio includes more than 50 programs, 600 courses, and nearly 1000 assessments.

Prior to joining WGU, DeMark spent more than 15 years at leading IT companies, serving in various leadership roles where she oversaw the strategy and execution of the design, development, and deployment of certification and curriculum-based assessment portfolios. Previously, she was an independent consultant working with state and local school districts, as well as working with The College Board on SAT and AP program evaluation.

DeMark is published in numerous journals and books and is a sought-after speaker. DeMark currently sits on ANSI’s Personnel Certification Accreditation Committee, which serves to validate whether certification programs adhere to standards.

DeMark earned a Ph.D. in Educational Psychology (Measurement, Statistics, & Methodological Studies) from Arizona State University. DeMark earned B.S. degrees in both Elementary Education and Psychology from Vanderbilt University.
Dr. Pradeep Kotamraju is currently the Bureau Chief, Career and Technical Education, Division of Community Colleges, Iowa Department of Education. As Iowa’s State Director for Career and Technical Education (CTE), he has leadership responsibility in managing those secondary and community college CTE programs that are funded through the Carl D. Perkins federal program. Previous to his current position as the Iowa CTE State Director, Dr. Pradeep Kotamraju has served the Deputy Director, National Research Center for Career and Technical Education (NRCCTE), University of Louisville, Louisville, Kentucky. Prior to that, he served as the System Director, Perkins, at the Minnesota State Colleges and Universities, Office of the Chancellor. Dr. Kotamraju has worked in several senior administrative positions in higher education and workforce development agencies in Minnesota.

Dr. Kotamraju has written several publications and monographs, and made numerous presentations, in the area of student success in career and technical education, workforce development in the United States, and, in the area of economic progress in the developing world. His research has included the examination of a variety of labor market information and workforce development issues that connect occupations, skills and careers, as individuals transitioned back and forth between employment and education. Dr. Kotamraju has been invited to participate on several statewide, regional and national committees that have focused on CTE programs, budget and finance, and accountability. Some of these committees have had even broader focus that places CTE right front and center when it comes to connecting education, workforce development, and economic development.

Before working in the public sector, Dr. Kotamraju taught college- and university-level Economics and Statistics at several higher education institutions in Minnesota and Kentucky. Dr. Kotamraju holds a Ph.D. in Economics from the University of Illinois. He received his Masters Degree in Economics from George Washington University, and his Bachelors in Economics from the University of Delhi, India.
Michael Morsches
Dean of Learning Enrichment and College Readiness
Moraine Valley Community College

Michael Morsches has worked in higher education for more than thirty years. His primary focus has been on developmental education and the transition from high school to college.

Michael currently serves as the Dean of Learning Enrichment and College Readiness at Moraine Valley Community College. He oversees the ABE/GED, ESL, developmental education, literacy volunteers, and tutoring programs. Michael has published numerous articles and handbooks on retention, student engagement, and teacher training in post-secondary institutions.
Yvette Mozie-Ross, Ph.D.
Vice Provost for Enrollment Management and Planning
University of Maryland, Baltimore County

Yvette Mozie-Ross, PhD, is Vice Provost for Enrollment Management and Planning at the University of Maryland, Baltimore County (UMBC). As Vice Provost, Dr. Mozie-Ross provides oversight and strategic planning for the areas of undergraduate admissions and orientation, financial aid and scholarships, academic and pre-professional advising, records and registration, and the student administration project (student information system). With a higher education career spanning over 25 years, she has served in numerous professional capacities including residence community director, coordinator of multicultural recruitment, assistant director for transfer recruitment and admissions, director of undergraduate admissions, and director of academic services (advising and registration). Dr. Mozie-Ross has served on various national and statewide committees and workgroups including the College Boards’ Commission for Transfer Policy and Practice, and the Maryland Higher Education Commission’s State Plan Writing Group on Access, Affordability and Completion. She has served on the university’s Strategic Planning Steering Committee and is currently serving as a member of the governing board for the Baltimore Collegetown Network, a consortium of 13 colleges in Baltimore, Maryland. Dr. Mozie-Ross frequently lends her expertise, both nationally and internationally, in the area of data analytics and leveraging analytics for institutional transformation. Dr. Mozie-Ross earned her bachelor’s degree from UMBC in 1988, her master’s degree from University of Maryland University College in 1994, and her doctorate in Education Policy and Leadership at the University of Maryland, College Park in 2011. Her dissertation research examined the academic and background characteristics of high school graduates who identified teachers as influential in their choice of college. Dr. Mozie-Ross enjoys spending time with her husband of 22 years and their 20-year old son. Her pass-time interests include family genealogical research and running.
Holly Zanville, Ph.D.
Senior Advisor for Credentialing and Workforce Development
at Lumina Foundation

Holly Zanville is Senior Advisor for Credentialing and Workforce Development at Lumina Foundation. She leads a new portfolio on Worker and Employer Engagement that focuses on building the capacity of educators and employers to scale and spread the best ideas in training, credentialing, and other workforce development strategies linked to postsecondary learning opportunities; and examining issues around the future of work and learning. Her work includes cultivation of networks and partnerships essential to the emerging new postsecondary learning system including Credential Engine, quality assurance efforts to ensure that credentials stand for high-quality learning, and networks for research and industry sector engagement. She previously led Lumina’s development of the national Connecting Credentials initiative, credential completion for returning adults with prior college/no credential, and statewide approaches to reverse-transfer degrees through the Credit When It’s Due initiative. Zanville received her Ph.D. in Educational Administration from the University of Minnesota; MA in English from the University of Wisconsin-Madison, and BA in English and Biology from Lindenwood University.
Appendix F. Expert Panel: Futurists
Notes of the Expert Panel Meeting Representing Futurists
June 21, 2018
National Assessment Governing Board
Ad Hoc Committee on Measures of Postsecondary Preparedness

As one step in addressing the charge of the Ad Hoc Committee on Measures of Postsecondary Preparedness, HumRRO organized and facilitated a meeting with a select group of futurists.39 The purpose of this meeting was to elicit input from thought leaders regarding the future of postsecondary education and work.

We were fortunate to assemble an exceptional panel of visionaries with a variety of perspectives. The panel members included Randy Bennett, Educational Testing Service; Karen Cator, Digital Promise; David Conley, EdImagine; Alana Dunagan, Clayton Christensen Institute; Devin Fidler, Rethinkery Labs, and Nancy Lue, Advanced Education Research and Development Fund on behalf of the Chan Zuckerberg Initiative and the Bill & Melinda Gates Foundation. Also, in attendance were several Governing Board members, Governing Board staff members, and HumRRO staff.

The meeting was held on June 21, 2018 in San Francisco, California. An overview of the National Assessment Governing Board and the charge of the Ad Hoc Committee on Measures of Postsecondary Preparedness, a “facebook” of attendees with brief biographic summaries, along with the agenda and logistical information for the meeting were sent to the panelists in advance of the meeting. Appendix A contains the agenda, list of attendees, and panelist biographies.

Terry Mazany, Ad Hoc Committee Chair, welcomed the futurists and set the stage for the role of NAEP in the future, given the impact of technology on work as well as the economic and global context in which students enter the postsecondary world. He led the attendees through introductions. Thanos Patelis (HumRRO) reviewed the agenda and stated the goals for the meeting.

To establish the perspectives of these varied experts, each panelist provided a 10-minute presentation of their initial thoughts regarding five discussion questions: (a) what are the trends you see that will define the future of learning and schools? (b) what are the trends you see that will define the future of work and the skills that will be most valued by employers of the future? (c) what are the most promising technologies that will redefine education? (d) what things are most likely to disrupt how we think about teaching and learning? and (e) what are the trends that most concern you, and why? Copies of the presentation slides are in Appendix B.

Following the presentations, Thanos Patelis facilitated deeper discussion about common themes and the five questions. Finally, Terry Mazany offered some concluding comments.

The purpose of this document is to summarize the themes and comments made by the panelists. The information in this report is meant to provide insight into the rich conversation and comments provided by the expert panelists.

---

39 Although some panelists would not describe themselves as “futurists,” per se, their careers all include the identification and evaluation of trends, as well as forecasting future conditions or developments.
Presentations

Randy Bennett described seven trends in the future of learning.

- Learning is increasingly technology-based with complex tasks (e.g., simulation and games).
- Materials and methods used in learning are only now catching up with cognitive science.
- Learning is more person-based, adaptive, and customized on different dimensions, to (a) allow accessibility to make learning more available to students with diverse learning types, (b) personalize in terms of competency level, (c) engage students effectively, and (d) give students greater agency over their learning goals.
- New constructs and competencies, such as socioemotional learning, citizenship and citizen engagement, and cross-cultural competency, are becoming more prevalent.
- Prior knowledge is critical when learning new information or developing new skills.
- There is a focus on cross-disciplinary skills such as communication and problem solving. However, contextual differences within disciplines are important considerations (e.g., problem solving in art differs from problem solving in science).
- Assessment embedded in instruction with automated analysis and feedback, allows for adjustment of instruction.

In addition to trends in the future of learning, Dr. Bennett described two trends of most concern.

- Personalization – There is concern that personalization could be used to exacerbate as much as ameliorate differences in opportunities and learning. For example, students from underrepresented groups could be routed toward basic skills classes.
- Embedding assessment in instruction – There is potential for embedded assessment in instruction for student learning, however conflating assessment for learning with assessment for accountability could be problematic, especially if used to make policy judgements.

Karen Cator provided the following perspectives regarding the five questions:

- Trends in the future of learning include: (a) personalization to accommodate variability in students through learning science, (b) more flexible learning to obtain and demonstrate competency, and (c) performance-based assessments leading to credentials for the changing global workforce.
- Trends in the future of work and skills include artificial intelligence (AI) which has the potential to disrupt many jobs. Employees will need deeper learning skills such as collaboration and social emotional skills. We should focus on what is uniquely human.
- Technology can be used to augment human performance. For example, data from embedded assessment and improved diagnostics can provide more precise and accurate analyses of student knowledge and performance, helping teachers perform more effectively in the classroom.
- Learning science could be disruptive. People will have jagged profiles—different levels of competence across skills—based on individual differences and the contexts in which they apply the skills.
- Most concerning is disenfranchisement of teachers. As an example, one-third of current teaching jobs in St. Louis are vacant. Other areas of concern include limited resources in

---

40 Ms. Cator recommended Jack Ma’s presentation at the World Economic Forum on The Way We Teach; https://www.youtube.com/watch?v=pQCF3PtaAsg.
schools, increasing cost of higher education, limitations of current assessments, equity of access to quality learning activities, and the digital learning gap.

David Conley shared the following insights regarding the five questions:

- The future of learning includes the following trends: (a) taking the teacher out of the bottleneck role, thereby allowing students to work at their own pace and receive just-in-time learning; (b) providing more social learning; (c) using technology to identify learning patterns to personalize learning; and (d) focusing on adapting skills to accommodate changes in work rather than learning fixed skill sets.
- Trends in the future of work and skills include changes such as (a) gig work versus long-term careers, (b) continued adaptability, (c) hiring at low- and high-skill end with less at the middle-skill level, (d) global work teams while living locally, (e) increasing service work, and (f) standardization versus bespoke work (see jagged profiles as mentioned by Ms. Cator).
- Promising technologies in education are adaptability, including a wider variety of students, specialized job/task-specific reading, and web-based learning.
- The following may contribute to disruptions in teaching and learning: (a) students having more agency over their learning, (b) basic skills taught in context using simulations or serious games such as used in the military and medical training, (c) self-directed learning will require resources for teachers to help students who have trouble directing their own work, and (d) emphasis on career preparation with certifications and badges over liberal arts education.
- The three most concerning trends are (a) equity in education, (b) equity in defining preparedness, and (c) increasing the pace of disruptive economic change.

Alana Dunagan offered the following insights:

- The future of learning and work includes the following trends:
  - increased online learning in higher education and K-12
  - learning not requiring a terminal degree (e.g., a certification)
  - workforce alignment of education
- Higher education institutions are seeing falling enrollment, while training in specific skills matters more. She described the parallels between disruptive innovations in education and in business. She explained that corporate bankruptcy following implementation of disruptive technology occurs when companies do not adapt by using technology to expand the reach of their services (i.e., they continue serving the same set of customers rather than expanding their customer base); Blockbuster is an example of this situation.
- Jobs requiring higher education are growing twice as fast as jobs that do not, because of disruption by the education technology market. Innovators in the education technology space are developing partnerships with employers and creating new ways of offering higher education providing the needed training.
- The biggest concern in education and work is the prestige-based model of signaling competence (i.e., a degree from an elite university is highly valued over a degree from a lower tier school without regard to a student’s actual knowledge and skill). This model ignores the skills a student has and does not include employers in identifying the skills that students should learn. A better model would engage businesses in identifying skill needs, offer education aligned to workforce needs, and provide students with evidence of skill attainment and a means for submitting that information to employers.
Devin Fidler provided a historical perspective to inform the following trends:

- The history of change in organization strategies evolved from guilds to industrialization to manufacturing/assembly to digital. The advent of the World Wide Web facilitated communication and has expanded to commerce and coordination.
- Examples of using technology to speed up work include peer to peer applications such as TaskRabbit, Gigwalk, and Upwork. These platforms have millions of people enrolled to offer their services with qualifications based on past performance. Employers can use these applications to identify well-qualified candidates with the appropriate skills mix and a history of positive reviews; employees can use these applications to find jobs and to see what skills are in demand.
- The most promising technologies are using organization technologies in education technology with artificial intelligence.
- Disruption will come from small innovative organizations who are more nimble than large businesses.
- The biggest concern is the stereotype that organization is dehumanizing; however, organization can expand human capability.

Nancy Lue identified the following education megatrends:

- Return on education (i.e., value of education) is similar to an internal rate of return (IRR), a term very popular in Silicon Valley to evaluate whether to invest in something. In 2015, 50% of college graduates were working in fields they did not study in school, and more than a third indicated they would study different subjects if they had the opportunity to do it all over again. In a 2013 survey, nearly 40% of college graduates indicated university did not prepare them for employment. Meanwhile, companies and non-profit organizations (e.g., Coursera, EdX, Khan Academy) offer courses for free or with a credential for $100. General Assembly boasts a 98% success rate in securing jobs or promotions within six months for their graduates.
- Continuous learning (e.g., Kaizen education) offers an opportunity for adults to keep up with education. Millennials are projected to have 15 careers in their lifetimes so this ongoing education is important. Coursera’s MOOC “Learning How to Learn” has been highly successful, with over half a million “alumni.”
- Technology provides opportunities for ongoing learning. Video games are one venue for learning. One way to get people interested in education is to make the best (aka, “rock star”) teachers available through technology.
- Knowledge increasingly can be seen as currency (e.g., micro-credentials, badges). The degree-driven, prestige-driven education system isn’t meeting the needs of modern society. “What you know” is becoming more important than “where did you go.” Individuals can curate a portfolio of skills evidenced by micro-credentials, etc.
- Big data and smart data provide an opportunity to use data to personalize learning (Dreambox, Knewton, etc.).
- Mobile technology learning applications are widely available. Today, 90% of high school and college students own a smart phone, and time spent on smart phones is increasing. Smart phones are providing opportunities to learn in small bits of time as well as in dedicated sessions.
- Mind, body, and soul: Breakthroughs in brain research and cognitive science are being incorporated into learning. Evidence is mounting that physical fitness, happiness, diet,
overall wellness, and mindfulness (e.g., Goldie Hawn’s MindUp program) are associated with successful learning.

- Equity is the greatest concern and pervades all of the trends. For example, education technology has costs which limit access. While mobile technology is available to many, ten percent of students do not have smartphones.

**Discussion**

Thanos Patelis (HumRRO) facilitated a deeper discussion among panelists about common themes and the discussion questions.

*Personalized learning.* Content can be tailored to student preparation, interest, and ability. Learning will feel more purposeful, connected, and relevant. Fewer students will be seated in rows in classrooms on a rigid schedule. In high school, students may enroll in work training programs or participate in micro-internships. Teachers will serve as mentors. There is a need to change the traditional school organization/culture and provide teachers with the knowledge and skills to educate students in a new environment.

*Contextual data.* Is a student goal-focused or not? Using data about students’ goals can improve instruction. Contextual data (e.g., goals, interests, self-confidence) may provide clues as to why a student might be struggling and may also provide insights to inform how to individualize instruction.

*Equity.* Opportunity to learn pervades multiple areas. Cost and availability can be barriers to access educational technology and higher education.

*Big data.* Educational technology generates a lot of data. Educators need to learn how to analyze and use the data, taking a data systems point of view. Also, there is a need to teach teachers how to capture and document performance data on what students are doing in the classroom and how to use those data to improve classroom instruction and activities.

*Data dashboards.* Data dashboards can connect data from different sources, interpret multiple data points, and provide evidence of what students can do (versus cannot do).

*Micro-credentials.* Micro-credentials can be used by students and teachers. Students could earn a micro-credential when mastering a concept. Teachers can use their students’ micro-credentials to identify the skills acquired and those that need to be taught or re-taught.

*Competency assessments.* Students would benefit from measures of job-related skills to show their potential and demonstrate performance capabilities, particularly if the measures do not correlate to student background. Employers benefit because they have evidence of a job candidate’s skills. Educators can use competency data to mentor students on achieving goals.
Panelist Recommendations
As a wrap-up exercise, Thanos Patelis asked each panelist to make one recommendation for the Governing Board to consider.

Randy Bennett – Use NAEP’s national probability sample to describe what instruction is like at different levels for different types of students (e.g., students with disabilities, socioeconomic status) across time.

Karen Cator – Work toward a more coherent assessment system across NAEP and states.

David Conley – Endorse the work of the Ad Hoc Committee with a longer-term vision for NAEP to be bold in creating better items and measuring traditional content with greater precision.

Alana Dunagan – Develop innovative methods to measure flexibility, problem solving, and non-traditional skills that people will need in the future.

Devin Fidler – Look at partnering with prestigious organizations within the learning space that function outside of formal assessment, such as skunk works and incubators.

Nancy Lue – Use NAEP to assess the technology gap and equity issue in technology use outside of the classroom.

Reflections
Terry Mazany expressed his appreciation for the panelists’ insights. He noted that each expert presented similar ideas through a different lens; while this might have seemed repetitive, it actually reinforced the conclusions. The panelists convinced him that traditional education enterprise is collapsing in slow motion. Innovation outside of education is occurring at an accelerating pace. Learning might occur in smaller units such as micro-credentials.

Mr. Mazany discussed the high cost of traditional higher education and the trillion-dollar impact of student debt on the economy. He acknowledged the existence of prestige-based signaling that maintains inequity in the system. These are complex and challenging social issues. NAEP may be able to be a market signal by Governing Board priorities regarding what to measure and report on. He opined that perhaps NAEP can reinforce that prestige alone is not the gold standard.
Appendix A: Meeting Agenda, Attendees, and Panelist Biographies

Futurist Expert Panel

Thursday, June 21, 2018 1:00 pm – 4:00 pm PT
Room: Cypress A * Hyatt Regency San Francisco Airport
1333 Bayshore Highway * Burlingame, California, USA, 94010

Agenda

1:00 – 1:15 pm  Welcome, Introductions, and Overview of the Ad Hoc Committee
Terry Mazany, Chair of the Ad Hoc Committee on Measures of Postsecondary Preparedness
Overview of the Agenda and Goals for the Meeting
Thanos Patelis, HumRRO

1:15 – 2:45 pm  Panelist Perspectives and Initial Thoughts Regarding the Discussion Questions
A series of ten-minute presentations, each followed by a five-minute Q&A.
1:15 – 1:30  Randy Bennett (Educational Testing Service)
1:30 – 1:45  Karen Cator (Digital Promise)
1:45 – 2:00  David Conley (EdImagine)
2:00 – 2:15  Alana Dunagan (Clayton Christensen Institute)
2:15 – 2:30  Devin Fidler (Rethinkery Labs)
2:30 – 2:45  Nancy Lue (Advanced Education Research & Development Fund)

Questions for Discussion:
1. What are the trends you see that will define the future of learning and schooling?
2. What are the trends you see that will define the future of work and the skills that will be most valued by employers of the future?
3. What are the most promising technologies that will redefine education?
4. What things are most likely to disrupt how we think about teaching and learning?
5. What are the trends that most concern you, and why?

2:45 – 3:45 pm  Panel Discussion
Facilitated by Thanos Patelis

3:45 – 4:00 pm  Final Reflections
Terry Mazany

Conducted in Support of the National Assessment Governing Board’s Ad Hoc Committee on Measures of Postsecondary Preparedness
Attendees

Expert Panelists:
- Randy Bennett, Norman G. Frederickson Chair in Assessment Innovation in the Research & Development Divisions, Educational Testing Service
- Karen Cator, President and CEO of Digital Promise
- David Conley, President, EdImagine
- Alana Dunagan, Researcher for Higher Education, Clayton Christensen Institute
- Devin Fidler, Founder, Rethinkery Labs
- Nancy Lue, Co-Lead, Advanced Education Research & Development Fund

Governing Board Members:
- James Geringer, former Governor of Wyoming
- Carol Jago, Associate Director, California Reading and Literature Project at UCLA
- Terry Mazany, Chair, Ad Hoc Committee on Measures of Postsecondary Preparedness
- Dale Nowlin, Teacher and Mathematics Department Chair, Bartholomew Consolidated School Corporation, Columbus, Indiana
- Linda Rosen, former Chief Executive Officer, Change the Equation, Washington, DC
- Chasidy White, Director of Strategic Initiatives, Office of the Superintendent, Montgomery, Alabama

Governing Board Staff Members:
- Michelle Blair, Assistant Director for Assessment Development
- Bill Bushaw, Executive Director
- Lisa Stooksberry, Deputy Executive Director
- Lily Clark, Assistant Director for Policy & Research

HumRRO Staff Members:
- Monica Gribben, Senior Staff Scientist
- Sunny Becker, Principal Staff Scientist
- Thanos Patelis, Principal Scientist
Ad Hoc Committee Meeting on Postsecondary Preparedness
Panelist Biographies

Randy E. Bennett Ph.D. ............................................................................................................49
Karen Cator ..............................................................................................................................50
David Conley, Ph.D. ..................................................................................................................51
Alana Dunagan .........................................................................................................................52
Devin Fidler ...............................................................................................................................53
Nancy Lue .................................................................................................................................54
Randy E. Bennett Ph.D.
Norman O. Frederiksen Chair in Assessment Innovation in the Research & Development Division Educational Testing Service

Randy E. Bennett is Norman O. Frederiksen Chair in Assessment Innovation in the Research & Development Division at Educational Testing Service in Princeton, New Jersey. Bennett's work has focused on integrating advances in cognitive science, technology, and educational measurement to create approaches to assessment that have positive impact on teaching and learning. From 1999 through 2005, he directed the NAEP Technology Based Assessment project, which included the first administration of computer-based performance assessments with nationally representative samples of school students, and the first use of "clickstream," or logfile, data in such samples to measure the processes used in problem solving. From 2007 to 2016, he directed an integrated research initiative titled, *Cognitively-Based Assessment of, for, and as Learning* (CBAL), which focused on creating theory-based summative and formative assessment intended to model good teaching and learning practice. Randy Bennett is president of the International Association for Educational Assessment (IAEA) (2016-), an organization primarily constituted of governmental and non-governmental nonprofit measurement organizations throughout the world, and immediate past president of the National Council on Measurement in Education (NCME) (2017-2018), whose members are individuals employed primarily in universities, testing organizations, state education departments, and school districts. He is a Fellow of the American Educational Research Association.
Karen Cator is President and CEO of Digital Promise and a leading voice for transforming American education through technology, innovation and research. From 2009-2013, Karen was Director of the Office of Educational Technology at the U.S. Department of Education, where she led the development of the 2010 National Education Technology Plan and focused the Office’s efforts on teacher and leader support. Prior to joining the department, Cator directed Apple’s leadership and advocacy efforts in education. In this role, she focused on the intersection of education policy and research, emerging technologies, and the reality faced by teachers, students and administrators. She began her education career in Alaska as a teacher, ultimately leading technology planning and implementation. She also served as Special Assistant for Telecommunications for the Governor of Alaska. Cator holds a master’s in school administration from the University of Oregon and received the 2014 College of Education Distinguished Alumni award. The American Association of Publishers has awarded Cator with the 2014 Visionary Award. She received her bachelor’s in early childhood education from Springfield College and received the 2015 Distinguished Alumna award. She is an Aspen Pahara Fellow, the past chair for the Partnership for 21st Century Skills and has served on boards including the Software & Information Industry Association-Education.
David Conley, Ph.D.
President, EdImagine
Professor of Educational Policy and Leadership in the College of Education at the University of Oregon
Director, Center for Educational Policy Research

David Conley is Professor of Educational Policy and Leadership in the College of Education at the University of Oregon where he directs the Center for Educational Policy Research. He is the founder and president of EdImagine, an educational strategy consulting company. Additionally, he founded and served for 12 years as CEO of the Educational Policy Improvement Center, EPIC (now Inflexion). He recently completed an appointment as Senior Fellow for Deeper Learning under the sponsorship of the Hewlett Foundation.

Dr. Conley is a national thought leader in the areas of college and career readiness, student ownership of learning, systems of assessment, educational accountability, and the future of education and the economy. He has published multiple articles and policy briefs as well as three books in these areas. His most current book, published by Harvard Education Press, is entitled *The Promise and Practice of Next Generation Assessment*.

He serves on numerous boards and advisory committees including as a member of the technical advisory committee of the Smarter Balanced Assessment Consortium (SBAC) and the Illinois State Board of Education Accountability Technical Advisory Committee, and as a founding board member of New Meridian, which now manages the PARCC assessments. Additionally, he chairs the New Meridian Steering Committee. Previously, he co-chaired the Validation Committee for the Common Core State Standards.

He has conducted multiple major research studies for the Association of American Universities, the College Board and its Advanced Placement program, the International Baccalaureate, and the National Assessment of Governing Board. He has most recently studied next generation systems of assessment, new indicators of college readiness, and new methods to determine career readiness.

Before entering higher education at the University of Oregon in 1989, Dr. Conley spent 20 years in the public-school system in a variety of roles including teacher and co-director of two alternative schools, a site and central-office administrator, and an executive in a state education agency. He is a first-generation college attendee who received his AA from Cabrillo College, his BA from the University of California, Berkeley, and his MA and PhD from the University of Colorado, Boulder. He grew up on the central coast of California, where he spent a great deal of time at the beach.
Alana Dunagan  
Researcher, Higher Education, Clayton Christensen Institute

Alana leads the Institute’s higher education research and works to find solutions for a more affordable system that better serves both students and employers. In this role, Alana analyzes disruptive forces changing the higher education landscape. Her research includes studying business model innovations, public policies, and investment strategies that can give rise to new and sustainable postsecondary models.

Prior to joining the Christensen Institute, Alana spent ten years in institutional investment management working on behalf of nonprofits, particularly colleges and universities. She worked as an investment consultant for Slocum, and spent five years with Macalester College managing their $700 million endowment. She holds a BA in Economics and Political Science from Macalester College and an MBA from the Harvard Business School.
Devin Fidler
Founder, Rethinkery Labs

Devin has worked with senior leaders at dozens of Fortune 1000 companies to systematically explore emerging issues and technologies, and to analyze their potential impacts. His ongoing work at Rethinkery Labs, including developing tools for “self-driving” management, has been covered by HBR, the New York Times, Wired and a number of other publications. He argues that today, companies themselves are a technology on the verge of disruption. Prior to founding Rethinkery, Devin founded and led the Future of Work and Future of Learning programs at the Palo Alto-based Institute for the Future.

Devin is a frequent speaker at gatherings of business leaders and others interested in the transformation of work and organizations. He approaches projects from a strongly international perspective, having lived and worked in several countries throughout his career.
Nancy Lue
Co-Lead, Advanced Education Research & Development Fund

Nancy Poon Lue is currently co-leading the exploration of a national Advanced Education Research & Development Fund on behalf of the Chan Zuckerberg Initiative and the Bill & Melinda Gates Foundation. She is also a Partner and Secretary of the Board of Directors of the venture philanthropy organization Silicon Valley Social Venture Fund (SV2). Previously, she served as Executive Director at the venture capital firm Global Silicon Valley (GSV) and was the inaugural General Manager of the EdTech Lab at GSVlabs. During the Obama Administration, Nancy was a Senior Advisor at the U.S. Department of Education where she led the development of the agency’s five-year strategic plan. Nancy is a Senior Fellow with the American Leadership Forum-Silicon Valley and sits on the Advisory Board of the AT&T Aspire Accelerator and the GreenLight Fund-Bay Area. She earned her B.A. and Ed.M. from Harvard College and Harvard Graduate School of Education.
Appendix B: Panelist Presentations

Note: These slides are the intellectual property of the presenters and should not be used or distributed for purposes beyond this Committee without permission.
Trends in the Future of Learning

- Be technology based, making greater use of complex tasks, games, simulations
- Be based on more modern underlying models of cognition and learning
- Be personalized in terms of:
  - Adaptability
  - Competency level
  - Background and interest
  - Learning goals
- Include (or give greater emphasis to) “new” competencies, e.g.:
  - Self-regulated learning
  - Citizenship, civic engagement
  - Cross-cultural competence
  - Using technology tools for problem solving
- Include traditional competencies:
  - Knowledge acquisition and construction
  - Made more (not less) important by technology
- Include focus on cross-cutting skills within the disciplines
  - Communication, critical thinking
- Embed assessment within instruction, including automated analysis and feedback

Thoughts on the Future of Education and Work

Randy Bennett
Educational Testing Service
Princeton, NJ 08541
rbennett@ets.org

Presentation as a member of the Futurist Expert Panel at the meeting of the National Assessment Governing Board's Ad Hoc Committee on Measures of Postsecondary Preparedness, San Francisco, CA, June 2018.

Trends in the Future of Work

- Continued automation of many types of manual, cognitive, and social-interactional work
- Pervasiveness of technological tools for problem solving:
  - Aids requiring constant proximal human interaction (Excel)
  - Extensions allowing for remote reach (drones)
  - Assistants: Carry out this subtask (Siri)
  - “Intelligent” implementers: Work independently with human QC

Overview

1. What are the trends you see that will define the future of learning and schooling?
2. What are the trends you see that will define the future of work and the skills that will be most valued by employers of the future?
3. What are the most promising technologies that will redefine education?
4. What things are most likely to disrupt how we think about teaching and learning?
5. What are the trends that most concern you, and why?
Most Likely Challenges to Disruptions for Teaching and Learning

- In K-12, the challenges are still greater than the disruptors
  - Level of, and extent of variation in, quality of teaching
  - Level of, and extent of variation in, school technology
  - Variation in funding for education by locale
  - Grade-based organization of schooling
  - Local control
    - little coherence, massive inefficiency due to no chance for economy of scale
  - Size (3rd largest country in the world)
  - Concerns for privacy of student data
  - Concerns over the corporatization of education
  - Public indifference; even antipathy, toward rigorous expectations and toward addressing inequality

Skills Most Valued by Employers

- Using technology tools for problem solving—i.e., to create value by being able to use:
  - Aids requiring constant proximal human interaction (Excel)
  - Extensions allowing for remote reach (drones)
  - Assistants: Carry out this subtask (3in)
  - “Intelligent” implementers: Work independently with human QC
- Being able to, individually and in collaboration with others, locate, evaluate, integrate, synthesize, apply, and construct knowledge (i.e., to learn)
- Being able to communicate, educate, and help others make effective decisions

Most Likely Disruptors

- People
  - What factors will make educators, policy makers, parents, students, and public advocate for, and accept, change?

Most Promising Technologies for Education

- Technologies that increase opportunities for remote social interaction
  - Learning is a social activity
- Adaptive learning (intelligent tutoring) combined with human instruction
- Simulations, games, virtual reality that pose tasks and situations similar to the ones students must learn to negotiate as proficient practitioners in a domain
- Analytics to help adapt instruction, guide students in managing their learning, help teachers improve instruction and its management
Trends of Most Concern

- Personalization
  - Equity: different foci of instruction by demographic group
- Idea of replacing end-of-unit assessment with embedded formative assessment
- Use of AI (without sufficient human oversight) for consequential decision-making purposes
  - Where explanation is important, current approaches to AI are insufficient for making decisions that affect life chances
  - EU GDPR requires provision of an explanation
1. What are the trends you see that will define the future of learning and schooling?

- Personalization
- Learner Variability (advancements in learning sciences)
- Competency based learning (+ performance assessment)
- World Challenges (e.g., UN SDGs)
- Workforce Changes
What are the trends you see that will define the future of work and the skills that will be most valued by employers of the future?

**Artificial Intelligence**
- Ability to learn
- Work with others
- Flexibility and comfort with complexity
- Creativity and solution development
- Compassional Thinking

What are the most promising technologies that will redefine education?

**Augment Human Performance**
- Data and instrumentation (like location & weather) - Adaptive
- Improved diagnostics and embedded assessment
- Moving from (average and comparison) to precision and accuracy
- Virtual and augmented reality
- Open Education Resources - organized, findable and contextualized
What things are most likely to disrupt how we think about teaching and learning?

- Learning Sciences
- Improvement Science
- Research and Evidence
- Advanced R&D - Pasteur's Quadrant

What we know - Learning Sciences

- Learning is Developmental
- Challenging & Achievable
- Total Environment
- Practice, Practice, Practice
- Personal & Meaningful
- Habits of Mind
- Prior Knowledge
- Emotion Matters
- Social Interaction
- Unique Jagged Profiles

Learner Variability

What are the trends that most concern you, and why?

- Lack of respect for and disenfranchised teachers
- Under-resourced schools
- Cost of higher education
- Assessments that fall far short of the full picture
- Issues of inequity
- Digital Learning Gap (Access - Participation - Powerful Use)
Conley Presentation
NAGB Presentation

the questions

1. What are the trends you see that will define the future of learning and schooling?
2. What are the trends you see that will define the future of work and the skills that will be most valued by employers in the future?
3. What are the most promising technologies that will redefine education?
4. What changes are most likely to change how we think about teaching and learning?
5. What are the trends that most concern you, and why?

Trends that will define the future of learning and schooling?
- Personalization of access to knowledge
- Universality of learning processes
- Involvement of learners in instruction or program design
- Use of stress learning
- Social learning
- Complex profiles of learners
- Importance of adaptability over fixed skills
- Need to focus on wider life competencies, adaptable learners

Trends that will define the future of work
- Day vs. career
- Low skill vs. high skill
- Place, time, set into
- Stronger, legacy business
- Globalization vs. localities
- Service vs product
- Automation/standardization vs customization/complexity
skills that will be most valued by employees of the future

- self-directed learning
-agneurology of health, skill mastery vs. social construct
- human learning and application of knowledge
- simulation, serious games
- self-directed learning
- design and design education
- modularization of learning
- reinforcement, feedback, experience

the most promising technologies that will redefine education

- innovation
- understanding the physiology of learning
- epistemology
- constructivism
- feedback
- interactive, stimulating environments
- learning environments that can prevent or cure academic gaps
- AI, big data, machine learning

things are most likely to disrupt how we think about teaching and learning

- self-directed learning
- emphasis on skills, skill mastery vs. social construct
- human learning and application of knowledge
- simulation, serious games
- self-directed learning
- design and design education
- modularization of learning
- reinforcement, feedback, experience
Appendix F. Expert Panel: Futurists
Dunagan Presentation
Disruption could create major benefits for students and society:

- Radically Affordable
- Accessible to All
- Knowledge + Skills
- Workforce Ready
- Motion + Structure
- Public + Private

Higher education is unaffordable
Today's Landscape: Enrollment

- Traditional and non-traditional students going to college:
  - Graph showing trends over time.

But our economy is more dependent on skilled workers than ever before.

Skills matter more than ever:

- Prepare students for jobs because they don't hire skilled workers.
- Jobs requiring skills:
  - Graph showing increasing demand or lack thereof.

Lots of new, innovative players emerging:

- Logos of educational and professional development providers.
1. Online learning is quickly becoming ubiquitous in higher education.

- Fall 2016: 30.8% of undergrads and 34.6% of graduate students are learning online.
- Online learning is especially prevalent among those who otherwise wouldn’t be able to access higher education.
- About half of students enrolled in school were learning online.

2. Personalization, competency-based education, student supports, and mentoring are important for experiential learning.

- College is still the currency of the labor market, but as Ryan Craig says, it’s a currency system with $100,000 bills and no smaller denominations.
- Automation and technology are changing the nature of work—and that change will be continuous.
- Learning won’t end with college; graduation...
- ...This creates big opportunities.
3) Workforce alignment built into program and curriculum design

- Building relevant certificates into programs
- Using industry experts (rather than academic experts) to design curriculum
- Creating explicit learn-to-work and learn-to-network opportunities
- Experiential learning

---

Keep the conversation going.

#disruptiveinnovation

---

[Logos and links]
Fidler Presentation
Much Bigger: Organizational Technologies

Preparing for Next Generation Organizational Technologies

Devon Miller | devon@thinkerslabs.com
There Are Many Ways to Organize

- 4500 years ago: First legal structures for companies
- 1500 years ago: Trade guilds encode industry across Europe
- 250 years ago: Industrial revolution and modern companies emerge
- 130 years ago: Assembly lines, globalization etc.
There Are Many Ways to Organize

- 1000 years ago: First legal structures for companies
- 150 years ago: Industrialization and modern companies emerge
- 20 years ago: Virtual reality, globalization etc.

Digital Organization is Special

- 200 years ago: First legal structures for companies
- 100 years ago: Trade policies as a model for Europe
- 50 years ago: Industrialization and modern companies emerge
- 10 years ago: Assembly lines, globalization etc.
Appendix F. Expert Panel: Futurists

Human Task Routing

Use emerging technologies to:
activate, deactivate & reconfigure resources
where they are needed & when they are needed

The Frontiers of “OrgTech”

A Collective Intelligence Engine

- Work that looks for people, instead of people looking for work
- Using platforms to find the best matches available
- Collective Intelligence Surfacing insights
Training, National Interest, and the New OS

Orchestration becomes a more powerful skill

Many more established industries can expect competitors built along these lines - Transition and national interest

Thank You!

Devor Adler | Devor@RethinkeryLabs.com

RETHINKERY LABS
Lue Presentation
National Assessment Governing Board Futurist Panel
June 21, 2018
Nancy Poon Lue

1. Return on Education (ROE)

gsx.com/2020-vision
Appendix G. Expert Panel: State Assessment Directors
Summary of the Expert Panel Meeting with State Education Officials
June 28, 2018

National Assessment Governing Board
Ad Hoc Committee on Measures of Postsecondary Preparedness

To support the charge of the Ad Hoc Committee on Measures of Postsecondary Preparedness, the Human Resources Research Organization (HumRRO) convened a small volunteer panel of education officials responsible for their state’s assessment and/or accountability. The meeting was conducted in partnership with the National Assessment Governing Board (Governing Board) and the Council of Chief State School Officers (CCSSO). The focus group was conducted on June 28, 2018, in San Diego, California during the CCSSO-sponsored National Conference on Student Assessment. The purpose of the focus group was to gather information about states’ definitions of postsecondary preparedness/readiness and their efforts to develop and use indicators of postsecondary preparedness/readiness.

The focus group participants included Chris Janzer, Michigan; Russell Keglovits, Nevada; Shelley Loving-Ryder, Virginia; Vaughn Rhudy, West Virginia; Michael Sibley, Alabama; Jenny Singh, California; Allison Timberlake, Georgia; and Vince Verges, Florida. Ms. Loving-Ryder and Mr. Sibley participated in the panel as both state experts and members of the State Policy Task Force, which is jointly convened by the Governing Board and CCSSO.

In attendance were Governing Board members Tyler Cramer and Joseph Willhoft; Governing Board staff members Michelle Blair, Lily Clark, Sharyn Rosenberg, and Lisa Stooksberry; CCSSO staff members Fen Chou and Scott Norton; and HumRRO staff members Sunny Becker, Monica Gribben, Thanos Patelis, Sheila Schultz, and Arthur Thacker.

An overview of the Governing Board and the charge of the Ad Hoc Committee on Measures of Postsecondary Preparedness, along with the agenda and the logistical information for the meeting, were sent to the participants as read-ahead materials. The meeting agenda is at Appendix A.

Thanos Patelis, HumRRO Principal Scientist, started the meeting by reviewing the agenda and goals. Lily Clark, Governing Board Assistant Director for Policy and Research, welcomed everyone and provided an overview of the Governing Board’s Strategic Vision initiative to “develop new approaches to measure the complex skills required for transition to postsecondary education and career,” which led to the creation of the Ad Hoc Committee on Measures of Postsecondary Preparedness and the impetus for this focus group meeting.

Mr. Patelis facilitated a discussion among the participants that highlighted the following guiding questions:

- How does your state define college and career readiness?
- Did your state consult with industry groups to define career readiness?
- What measure(s) does your state use to assess career readiness?
- Is military service a component of postsecondary readiness in your state?
- How does your state use noncognitive measures?
• Are there innovative or nontraditional indicators that your state might use to measure or report on students’ college and/or career readiness (e.g., student interest, micro-credentials earned, work-based learning)?
• What NAEP reporting on postsecondary readiness would be useful to states?

Following is a general summary of the information provided by this group of state assessment and accountability experts on definitions, activities, and indicators of postsecondary preparedness/readiness.

Definitions
The state officials offered examples of definitions of college and career preparedness/readiness used in their respective states. It was evident from the examples that states have a variety of definitions for college and career readiness. The definitions and indicators for college readiness were separate from those of career readiness. Most of the definitions for career readiness explicitly included “soft skills,” such as communication, collaboration, problem solving, and business practices. The state officials acknowledged the importance of soft skills to college and career readiness while also noting the challenge they pose in developing and measuring indicators related to these skills.

The definitions of college and career preparedness/readiness represented by the participating state officials varied in certain aspects and included the following:

- Two states defined **college readiness** as students who enroll and succeed in college courses without remediation.
- The use of benchmarks on college entrance and placement tests serve as a default definition of **college readiness**.
- **Career readiness** can be defined as obtaining a job that pays a living wage, which varies by location.
- **Career readiness** in several states was defined by a set of credentials from a career and technical education (CTE) program that did not include inter- and intra-personal skills. However, some other states included soft skills, such as inter- and intra-personal skills and business skills, in their definitions.
  - In one state, the inclusion of service learning was part of the secondary school experience that contributed to a career ready diploma seal.
  - In another state, career readiness was defined as acquiring specific skills from CTE programs as well as successful performance on assessments that represented specific skills (e.g., National Occupational Competency Testing Institute) and experience in a simulated workplace program.
- One state described the development of **college and career readiness** standards that defined specifically what is meant by college attendance and students’ understanding of the available career fields.
- **Military readiness** was offered as a postsecondary option that involves a set of cognitive and physical requirements, which is seen as an indicator of readiness in some state accountability plans.

A couple of state officials commented how they would welcome a definition of college and career readiness from the Governing Board.
Learning Opportunities and Interventions
Several state officials described the following efforts for students to acquire college and career readiness skills:

- States work with schools and industry to develop diplomas to certify technical career skills.
  - The diploma is earned through CTE programs, work-based learning, industry/credential exams, or portfolios.
  - One state developed career ready diploma seals that reflect cooperation between CTE programs and industry to introduce service learning and experiences for students to acquire industry-specific technical and broad interpersonal and intra-personal skills (e.g., leadership, collaboration, communication skills).

- Programs to prepare students for career readiness are designed to take advantage of local industry and involve the cooperation and input of businesses likely to hire postsecondary students.

- Schools encourage or adopt dual enrollment initiatives to increase student access to college-level courses and experiences.

- Soft skills, such as communication and leadership skills, are taught through service learning, student organizations, work-based learning, and simulated work environments.

- One state’s goal is to prepare students for college or a career by ensuring they are agile in facing an environment where the requirements are not always known.

- One state official indicated that the state department of education is (and should be) flexible in facilitating local education agencies to develop pathways for students that are relevant for local conditions and situations.
  - As an example, one school district described a multi-national company that moved into the municipality with plans to add an international business pathway for students. Students who complete designated international business courses and activities earn a career ready seal on their diplomas.

Data and Indicators
The state officials identified sets of skills important for college and career readiness. Some commented on the difficulty in measuring certain skills from both practical/logistical and technical/measurement perspectives. One state official opined that it is easier to measure college readiness than career readiness. Many state officials noted the difficulty with career readiness data is twofold: (a) the skills to be assessed are multi-faceted in nature and (b) there are practical limitations in identifying measurable indicators for all facets.

The skills explicitly mentioned, especially for career readiness, include business practices, collaboration, leadership, communication, creative problem solving, argument and reasoning, designing solutions, time management, and intellectual curiosity.

Several state officials indicated the Governing Board could contribute to the measurement of the soft skills important for indicating career readiness, particularly if provided at the state level. One official, however, encouraged the measurement of both college and career skills, but also cautioned that one consequence of reporting these skills by state is how industry may use them to target or avoid certain states for opening corporate and business locations.

State officials offered various comments and suggestions about data related to college and career readiness:
• Geographic differences reported in relevant career skills were based on the types of local industry and available jobs. States want data at a regional level.
• Some soft skills are not easily defined or measured (e.g., time management, intellectual curiosity).
• Student level data on absences, credits, and required course attainment can serve as proxies for some soft skills.
• A portfolio of artifacts (in the form of certificates, work-based learning, etc.) or experiences (advanced courses, dual credit) can be used as an indicator of college and career readiness.
• A concern about equity in terms of (a) opportunities to learn and (b) distribution of funds to offer college and career readiness opportunities (test fees) was expressed.
• Student service learning could be used as a relevant data point.
• One suggestion was for states to support and incorporate local accountability plans and metrics that involve school-specific indicators of important constructs such school culture, climate, and other environmental measures.
  o Examples of using school climate and school culture surveys were reported.
• Indicators used in state accountability plans included attendance, course participation, college entrance and placement test scores, and certification test results.

Various comments were offered about the measurement of college and career readiness:

• College readiness is easier to measure than career readiness.
• Soft skills typically are not included in state standards, so what to measure becomes a challenge.
• Measures should be general (versus specific) to remain relevant over time.
• Soft skills should be measured early (e.g., age appropriate elementary and middle school skills) to allow time for students to close gaps and attain common school and workplace skills. Early measurement would provide schools with data to monitor student learning and acquisition of these important life skills.
• States would like to see best practices in providing, documenting, and measuring college and career readiness skills.
  o For example, is there evidence that students who earn certificates are successful?
• A couple of state officials commented that the Governing Board is in a unique position to develop a measure(s) of soft skills at the state/national level.
• It would be a tremendous contribution if the Governing Board created a single definition inclusive of both college and career readiness as well as developed indicators to measure those skills.
Appendix A: Meeting Agenda and Attendees
Discussion of State Efforts on College and Career Readiness

Thursday, June 28, 2018, 7:30 – 8:50 AM PST
Room: Cobalt 520 (Level 5)
Hilton San Diego Bayfront
San Diego, California

Agenda

Purpose: Identify and discuss states’ current and innovative practices regarding college and career readiness to inform the National Assessment Governing Board’s effort to “Develop new approaches to measure the complex skills required for transition to postsecondary education and career.”

7:30 – 7:45 AM Breakfast & Introductions

7:45 – 8:00 AM Overview of the National Assessment Governing Board’s Initiative on Postsecondary Preparedness
Lily Clark, Assistant Director for Policy and Research
National Assessment Governing Board

8:00 – 8:50 AM Discussion of State Efforts on College and Career Readiness
Thanos Patelis, Facilitator, HumRRO

Guiding Questions:

- How does your state define college and career readiness?
- Did your state consult with industry groups to define career readiness?
- What measures does your state use to assess career readiness?
- Is military service a component of postsecondary readiness in your state?
- How does your state use non-cognitive measures?
- Are there innovative or non-traditional indicators that your state might use to measure or report on students’ college and/or career readiness (e.g., student interest, micro-credentials earned, work-based learning)?
- What NAEP reporting on postsecondary readiness would be useful to states?

8:50 AM Thank you and Adjourn
Attendees

State Officials (Department of Education)
Chris Janzer, Michigan
Russell Keglovits, Nevada
Shelley Loving-Ryder, Virginia
Vaughn Rhudy, West Virginia
Michael Sibley, Alabama
Jenny Singh, California
Allison Timberlake, Georgia
Vince Verges, Florida

CCSSO Staff Members
Fen Chou
Scott Norton

National Assessment Governing Board Members
Tyler Cramer
Joe Willhoft

National Assessment Governing Board Staff Members
Michelle Blair
Lily Clark
Sharyn Rosenberg
Lisa Stooksberry

HumRRO Staff Members
Sunny Becker
Monica Gribben
Thanos Patelis
Sheila Schultz
Arthur Thacker
The summary of the Young Adult Expert Panel meeting is forthcoming.
Appendix I. Literature Review: Work of the Future
Work of the Future – 2030

Overview of Jobs of the Future

History shows major changes in the occupational landscape and pace of life with each of the four industrial revolutions (Vale, 2016). The first industrial revolution, characterized by the steam engine, led to greater dispersal of jobs as those requiring machine power were not restricted to locations with wind or flowing water to power mills. Electricity and mass production brought about the second industrial revolution, leading to a surge in manufacturing jobs and supporting industries such as transportation, sales, and business. The advent of the digital age, the third industrial revolution, gave us the ability to collect and process massive amounts of data quickly and opened up new jobs related to computers and technology innovation. Now, we are entering the fourth industrial revolution, highlighted by the internet of things and artificial intelligence (Choi, 2017; Vale, 2016).

Throughout history, the introduction of new technologies has led to changes in jobs, from replacing workers with machines to changing how people perform their job to creating new occupations. According to futurists, this trend will continue. Policy analysts predict up to 47 percent of jobs in the United States could be automated between 2017–2037 (Bakhshi, Downing, Osborne, & Schneider, 2017; Houser, 2017). Opportunities will become limited in many industries, mostly in low- or medium-skill jobs, as automation reduces the number of humans needed to perform routine tasks. Further, business leaders and strategists predict that 50 percent of the occupations of 2014 will no longer exist in 2025 (Andrew, Ip, & Worthington, 2014). Technology, automation, artificial intelligence, and other innovations that have yet to be developed will lead to new occupations and jobs.

Atkinson and Wu (2017) take a different perspective of technological disruption, suggesting that others have based their doomsday predictions of rampant job loss on “faulty logic and erroneous empirical analysis.” Instead, they calculate, from 2010 to 2015, approximately six technology-related jobs were created for every 10 lost, the lowest share of jobs lost to technology of any period since 1950 to 1960.

While there will likely be changes in jobs and occupations of the future, what those changes will be, the extent and pace of changes, and the impact on employees entering or currently in the workforce are equivocal. In this report, we review the research related to potential changes in the workplace and highlight forecasts of jobs of the future.

Projections of Shifts in Jobs

Prediction of widespread unemployment due to technological advances is nothing new. For example, in the 1930’s, John Maynard Keynes predicted large-scale job loss associated with new technologies (as cited in Frey & Osborne, 2013). Recently in the United States, automation has been replacing jobs faster than it can create them (Atkinson & Wu, 2017; Brynjolfsson & McAfee, 2011), although perhaps not as quickly as some suggest. Autor, Levy, and Murnane (2003) found that as industries use automated technology to reduce the cost of performing routine cognitive and manual tasks, they hire more people to perform nonroutine cognitive tasks.

The occupations in which people are or will be employed are expected to shift, but this does not necessarily mean current jobs will be totally eliminated. As Manyika (2017a) reports, at least 30 percent of activities for most occupations could be automated using current technology. Assuming in many current occupations certain activities or tasks will be automated, current jobs
will change and more people will need to work with technology. Although some employees may lose their jobs because automation will drastically eliminate the need for human skills, integration of technology will help other workers perform their job better or enable them to be more efficient or productive. For still other workers, the demand for their skills may increase or the nature of what they do and how they accomplish tasks at work will change.

**O*NET Projections**

O*NET OnLine (National Center for O*NET Development, 2018) is a rich source of “detailed descriptions of the world of work.” There is a wealth of data available to those looking for work or interested in changing careers, as well as support for workforce development and human resources professionals, researchers, and policy analysts.

Using 2016–2026 employment projections from the Bureau of Labor Statistics, O*NET includes a set of Bright Outlook occupations. Twelve of the Bright Outlook occupation categories (including 20 distinct occupations) are expected to grow rapidly with an employment increase of 10% or more and are forecasted to have 100,000 or more job openings between 2016 and 2026 (see Table 1). O*NET identifies occupations linked to the green economy, focused on reducing environmental risks and initiating sustainable development without degrading the environment. Green jobs identified in O*NET are those where changes are expected in job demand, including work requirements such as tasks performed or worker qualifications such as knowledge, skills, and credentials needed for employment in these positions.

**Table 1. O*NET Bright Outlook Occupations with Rapid Growth and Numerous Job Openings**

<table>
<thead>
<tr>
<th>Major Occupation Group</th>
<th>Occupation Category</th>
<th>Green Economy Sector</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business and Financial Operations</td>
<td>Accountants and Auditors</td>
<td></td>
</tr>
<tr>
<td>Education, Training, and Library</td>
<td>Teachers and Instructors, All Other (includes Tutors)</td>
<td></td>
</tr>
<tr>
<td>Healthcare Practitioners</td>
<td>Registered Nurses (includes Acute Care Nurses, Advanced Practice Psychiatric Nurses, Critical Care Nurses, and Clinical Nurse Specialists)</td>
<td></td>
</tr>
<tr>
<td>Healthcare Support</td>
<td>Home Health Aides</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nursing Assistants</td>
<td></td>
</tr>
<tr>
<td>Personal Care and Service</td>
<td>Personal Care Aides</td>
<td></td>
</tr>
<tr>
<td>Food Preparation and Serving</td>
<td>Cooks, Restaurant</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Combined Food Preparation and Servicing Workers, Including Fast Food</td>
<td></td>
</tr>
<tr>
<td>Building and Grounds Cleaning and Maintenance</td>
<td>Janitors and Cleaners, Except Maids and Housekeeping Cleaners</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Landscaping and Groundskeeping Workers</td>
<td></td>
</tr>
<tr>
<td>Sales</td>
<td>Sales Representatives, Services, All Other (includes Energy Brokers)</td>
<td>X</td>
</tr>
<tr>
<td>Construction and Extraction</td>
<td>Construction Laborers</td>
<td>X</td>
</tr>
</tbody>
</table>

Source: O*NET OnLine
Several researchers have mined the O*NET data to make predictions about the future of jobs, identifying ones expected to increase and ones to decrease in the future. Bakhshi et al. (2017) used O*NET’s importance ratings in foresight exercises to generate input for a machine learning model, with the goal of mapping O*NET knowledge, skills, and abilities variables to future occupational demands. In the United States, the model predicts increased demand for teachers from pre-school through high school and post-secondary. Animal care workers, legal professionals, and engineers round out the top five occupations with expected increased demand.

**Jobs Expected to be Lost or to Decrease**

Job loss and decrease due to technology is evident all around us. Grocery stores offer multiple lanes where customers scan their own purchases, monitored by a single cashier. Only a few years ago there were multiple cashiers serving the customers. Financial institutions offer more and more automated functions such that their customers need to interface with a person less often than ever before. For example, customers can now use a mobile application to deposit a physical check without leaving home. Andrew, Ip, and Worthington (2014) predict customer work will disappear and many middle management positions will no longer exist in 2025. Frey and Osborne (2013) describe greater use of data and algorithms to computerize cognitive tasks such as fraud detection, health care diagnostics, legal document review, and financial advice. With automation expanding into more cognitively-advanced occupations, demand for individuals with certain professional skills, such as financial analysts and law clerks, are predicted to decline.

Frey and Osborne (2013) used O*NET data to study how susceptible jobs are to computerization. Using data for 702 occupations, they modeled the potential for jobs to be automated within 10–20 years. Their model predicted workers are most likely to be replaced with technology in occupations that involve transportation and logistics, office and administrative support, manufacturing, and service.

Bakhshi et al. (2017) did not use their model to predict decreased demand, but rather to predict low probability of increased demand. Those in skilled and semi-skilled trades, such as woodworkers, printing workers, metal and plastic workers, and other production occupations, were at the top of the future low demand list. Financial clerks received low ratings comparable to those in the trades.

**Jobs Expected to Increase**

Bakhshi et al. (2017) expect growth in professional occupations that require creative, digital, design, and engineering expertise. In addition to creativity, Osborne and Frey (n.d.) suggest growth in jobs that require social intelligence and manipulation, such as iOS and Android developers, social media interns, big data architects, data scientists, user interface/user experience (UI/UX) designers, Zumba instructors, and beachbody coaches. Further, strong interest in environmental sustainability is expected to benefit individuals employed in architectural and green occupations (Bakhshi et al.). Also, they foresee increased roles for people specializing in work reorganization, such as management analysts and training specialists.

**New Jobs to be Created**

Jobs requiring creative intelligence and social and emotional intelligence are predicted to be added to the economy, as are positions requiring the ability to leverage artificial intelligence (AI; Andrew, Ip, & Worthington, 2014). New jobs using creative or social and emotional intelligence or AI are expected to be more fulfilling than current jobs.
Generally, specific details about future jobs are scarce. Wagner (2011) discusses 70 jobs likely to exist in 2030. These jobs will be created through (a) retrofitting or adding new skills to existing jobs, (b) blending or combining functions from different jobs or industries, and (c) problem solving or creating new jobs to solve a problem. Types of jobs that might be added through retrofitting could support commercial space travel, such as space construction, space suit repair, space junk recyclers, astro-teachers, and exozooologists. By blending careers, the future might include environmental health nursing to treat patients exposed to toxins. To provide authoritative news in an era when anyone can publish online may lead to authority-journalists who specialize in an occupation and are cross-trained to report about their field. To solve future problems, we may hire digital footprint managers or digital archaeologists or future-guides. Wagner mentions occupations in the sustainability and green energy industries may be added, such as green career coach, autonomous vehicle operator, energy harvester, drone dispatcher, smart car interior designer, smart road designer/engineer, and smart road sensor control monitor/analyst. Gordon (2011) predicts there will be new careers inspired by nanotechnology, such as bio-botic physicians and bio-botist assistants to integrate biological functionalities and implanted nano-robotics to extend life.

Drivers of Change

The literature discusses three major drivers of projected shifts in jobs of the future—technology (Frey & Osborne, 2013), artificial intelligence (Manyika, 2017b), and social changes (Bakhshi et al., 2017; Manyika, 2017a). With changes in jobs come adjustments in the workplace. Experts predict that workplace culture and processes will shift as well as career paths and how people learn the necessary job skills needed to perform jobs of the future.

Impact of Technology

Literature is replete with observations of the accelerating impacts of technology in recent decades, including predictions this acceleration will continue. Baby boomers remember a world when communication required a phone call that was timed when both parties were available to speak or a letter that took days to be delivered; researching a topic involved going to the library or referencing a home copy of encyclopedia volumes; and getting a flat tire meant a hike to find a pay phone. Today’s young people are digital natives. They cannot imagine a world before e-mail allowed asynchronous communication; the internet offered a wealth of instantaneous information at one’s fingertips; and cell phones connected individuals to worlds beyond measure. The explosion of technology is expanding in multiple directions—and quickly.

Bakhshi et al. (2017) employed an innovative approach to predicting job trends by first paneling experts in “foresight workshops” and then inputting their expert judgments into a machine learning system. Their analysis of the experts’ judgments identified three key trends in technological change. First, fears about the impact of automation on employment are enduring. Second, estimates of the impact of future automation range from 9–47 percent of U.S. employment. Third, technology can amplify human performance and bring about new occupations and sectors.

For its 21st annual survey of CEOs worldwide, PwC interviewed 1,293 CEOs in 85 countries, including 104 from the United States, in October and November 2017 (Ryan, Sapin, Rao, & Ampil, 2018).

Based on these interviews, U.S. CEOs were hiring for broadly relevant digital skills and collaborative, creative, and efficient work styles. About two-thirds (63%) of those who were hiring found it more difficult to identify qualified workers than before. Responses to this same
survey indicated that artificial intelligence (AI) will be the innovation of the next two decades. CEOs predict that many workers will need AI literacy.

One of the challenges of a rapidly changing work environment is the ability of workers to keep pace. While new entrants into the workforce will grow up and attend school immersed in state-of-the-art technology, continued innovation ensures even these digital natives—those who have an advantage over older, digital immigrants who completed school before digital technology became omnipresent—will require ongoing training to stay current on technological knowledge, awareness, and skills. Employees who joined the job market prior to many of the current technological advances are already challenged with staying up to date. Two approaches to keeping tenured employees abreast of the latest technological developments are upskilling and reskilling.

**Upskilling**

When an employee upskills, that individual learns new skills to improve performance on the job or to adapt to new requirements of the job. Upskilling has the advantage of retaining experienced employees, a positive outcome as these employees are a known commodity to the employer, absent the risks of employing a new hire who may not be a good fit. Knowledge of corporate procedures, norms, and expectations eliminate the need for orientation and start-up time, and reduce the probability of missteps. Retaining seasoned employees also supports the maintenance of institutional memory, which can be crucial as an organization evolves and grows.

Training to upskill employees may be sought by the employee, imposed by the employer, or both. The PwC’s Workplace of the Future study found three-quarters of respondents expressed willingness to update their own skills. At the same time, most responding CEOs acknowledged an ongoing responsibility to upskill their employees (Ryan et al., 2018).

PwC’s Annual Global CEO Survey specifically investigated the employer’s perspective on upskilling. Nearly two-fifths (39%) of respondents reported initiating or using continuous learning initiatives to provide development paths for employees to gain skills.

**Reskilling**

When an occupation becomes obsolete or the changing nature of the position no longer suits an employee, reskilling may be in order. More disruptive than upskilling, reskilling is training an employee to perform an entirely different job.

Results from PwC’s annual survey of CEOs indicate companies that “reinvent their own talent” by reskilling their employees will have an edge by creating pathways for employees to better contribute to data-driven initiatives, which may lower costs and improve the customer experience among other impacts (Ryan et al., 2018). However, the U.S. lags other large economies (e.g., Germany, China, Japan) in assuming responsibility for retraining after automation (i.e., robots and AI) has been introduced into a job. The authors conclude that, while automation will result in job losses, over time those will be generally offset by new jobs. They note that “retraining workers to work with the support of AI will be important to future economic success” (Ryan et al., 2018, p. 15)

**Working with Data**

Alec Ross, author of Industries of the Future, provides an historical perspective of the workplace. He describes land as the raw materials of the agricultural age, followed by iron in the industrial age, and data in the information age. He posits that whoever owns, controls, and/or
can harvest meaning from data will define the future workplace. Ross (2016) emphasizes the sheer quantity of data being produced in recent history and the opportunity for data analytics to mine those data. For example, he notes that “90 percent of the world’s digital data has been generated over the last two years” (page 154). He opines that the sum of “all data from paintings on cave walls through 2003, we now produce every two days” (Ross, 2017).

PwC issued a report on the workforce of the future, using findings from a survey of 10,000 people in China, India, Germany, the United Kingdom, and the United States on how they think the workplace will evolve. From the survey findings, the authors developed four “Worlds of Work” for 2030 to describe hypothetical future scenarios defined along two continuums: collectivism and fragmentation. Authors concluded the increasing use of digital platforms and AI mean data are key. With augmented intelligence, humans and machines collaborate to make decisions. Uniquely human traits of emotional intelligence, creativity, persuasion, and innovation become more valuable. Adaptability will become increasingly important as work changes (PwC, 2017).

**Human-Technology Interactions**

Not only are data produced and stored at astounding rates, but individual access to such data through technology is expanding. Ross (2017) asserted that 20 billion networked devices were in circulation in 2017. He projected this number will reach 45 billion in 2020. This growth will likely not produce a steady expansion across all markets, but rather result in bursts of growth in traditional areas that have not been as impacted by the digital economy, such as transportation or mining.

PwC’s Annual Global CEO Survey predicts that businesses will initiate upskilling initiatives to teach employees the skills they need to augment their own work with the support of technology. The authors contend that companies will infuse AI into all aspects of their business, not just technology-related areas (Ryan et al., 2018).

**New Technology Jobs**

Technology jobs such as software engineers are on the rise, but two other trends may result in new technology jobs. First, the blending of AI technology with a human component, or augmented intelligence, may open opportunities for technology-enhanced versions of jobs that are available today (PwC, 2017).

Second, Ross (2016) points out an increasingly popular conviction that the opportunities of the future will no longer rigidly distinguish technical fields from liberal arts or humanities. He suggests hybrid studies will become more prevalent, such as a combination historian/electrical engineer or political scientist/computer scientist. He describes the thinking of Toomas Ives, President of Estonia: “…domains previously occupied only by people with backgrounds in the liberal arts, like government, will become increasingly occupied by people with more background knowledge in science and technology” (page 246).

**Impact of Artificial Intelligence**

**Types of Artificial Intelligence**

Although people may mean different things when they refer to artificial intelligence (AI), they generally mean the use of computers to perform tasks that require cognition and learning without programming the steps of the task. Often, AI is used to refer to machine learning, “where computers are taught or self learn how to recognize things” (Shaw, 2017). Bughin et al. (2017) describe other types of AI, including computer vision, autonomous vehicles, natural language, smart robotics, and virtual agents.
Machine learning is intensive, for the humans who must provide the “training data” and for the computer to process the information. Shaw notes that machine learning has many applications, such as predicting nefarious behavior or mechanical breakdown and identifying possible disease in 3D radiology images. Research is underway to explore the use of AI to make machine learning more efficient and accurate.

Shaw expects computer vision, using cameras to infer what they are seeing, to become the most prevalent type of sensor. Computer vision will be integral for self-driving cars and other autonomous vehicles such as self-driving trucks, buses, trains, and ships. Autonomous flying drones, which may be used for package delivery or to aid in aerial search and rescue, also will benefit from computer vision.

Natural language processors are familiar to many as they ask Siri for directions or to settle a debate. Smart home devices such as Google Home or Alexa are natural language processors. Once these devices understand what a person has said or written, Shaw states that a virtual agent is the next step. The virtual agent can help the human, provide financial advice, perform basic health diagnosis, or guide an individual through steps of an activity or job. Smart robotics are in use today, especially in manufacturing. Shaw expects robotics to become more prevalent in medicine, cleaning, stocking, agriculture, and food service in the future.

Machine learning is but one way of many to categorize AI. Hintze (2016) defines AI using a hierarchy from type I-reactive machines (e.g., Deep Blue, IBM’s chess supercomputer) to type II-limited memory (e.g., self-driving cars monitor information over time) to type III-theory of mind (e.g., understanding that thoughts and emotions affect behavior) to type IV-self-awareness (e.g., being aware of oneself).

**New Artificial Intelligence Jobs**

Research continues to advance AI (Bughin et al., 2017; Hintze, 2016; Shaw, 2017), with the implication that jobs developing and studying AI will continue to grow as the technology is incorporated into more daily life routines. Research firm Gartner, as cited in Singh (2017), predicts by 2020 more jobs will be created by the expansion of AI than will be lost. They estimate, that although AI will be responsible for the loss of 1.8 million jobs between 2018 and 2020, AI will create 2.3 million jobs. Healthcare, the public sector, and education will lead the way in incorporating AI into their sectors.

Increasing use of digital platforms and AI mean data will be key to creating new AI jobs (PwC, 2017). With augmented intelligence, humans and machines must collaborate to make decisions. Singh (2017) expects one in five workers will rely on AI to assist them in their jobs by 2022. It will be important for people to learn to work with and alongside AI machines.

**Impact of Social Changes**

**Globalization**

Globalization refers to the increasing interconnectedness of the world, both economically and politically. Along with automation, globalization is viewed as one of the main factors shaping the future workforce (Bernstein, 2016; Simon, 2016).

Companies operating on an international scale may have financial incentive to move jobs from the U.S. to other countries. This has been demonstrated historically through the loss of low-skilled manufacturing jobs due to offshoring (Hatzichronoglou, 2005). Today, higher skilled jobs also run the risk of offshoring, including computer-oriented science, technology, engineering, and mathematics (STEM) jobs (Lim, 2016).
Technological advances are closely linked with globalization’s impacts on the workforce. “Labor-linking” technology allows geographically dispersed people to vie for the same job, creating a competitive environment that could potentially drive wages down (Basu, 2016). However, there remain numerous higher-skilled jobs that are less subject to this threat, particularly those in healthcare and service industries that require face-to-face interactions (Blinder, 2007).

The potential for interaction with clients, customers, and coworkers from across the globe has implications for the skills that are valuable as well as valued. For example, employees may find it easier to negotiate the workplace when they have the skills needed to communicate effectively with geographically dispersed people from a range of sociocultural backgrounds. Employers are increasingly recognizing the value of cultural competence and communication skills among new hires (Vozza, 2016), especially when those skills are needed to perform future jobs that involve interaction on a global scale.

**Environmental Sustainability**

Environmental policies have long been linked to the reduction of jobs in specific industries (e.g., coal), though a causal link is up for debate (Morgenstern, Pizer, and Shih, 2001). The availability of jobs in such industries in the U.S. may in fact be limited by a decreased demand for fossil fuels that has resulted from advances in energy efficiency technology (Magill, 2017). Regardless of the mechanism at work, there is reason to believe the jobs of the future will continue to be shaped by both policy and consumer behavior related to environmental sustainability.

Beyond contributing to the obsolescence of some jobs, the focus on environmental sustainability continues to create new jobs and to change jobs that currently exist. Job opportunities for innovation related to environmental sustainability may increase as individuals and corporations alike seek to reduce energy consumption and waste (Bakhshi, Downing, Osborne, & Schneider, 2017). Companies that are changing practices to reduce their environmental footprint may create jobs for “sustainability professionals” who will take on the role of managing company resources (Hamilton, 2012). The National Center for O*NET Development has identified green economic sectors, green increased demand occupations, green enhanced skills occupations, and green new and emerging (N&E) occupations, many of which will likely boast increased job opportunities in coming years. Green enhanced skill jobs are those in the existing occupation that require significant changes due to the impact of the increased focus on environmental sustainability (O*NET, 2018).

**Demographic and Population Patterns**

As of 2016, foreign-born workers constituted nearly 17% of the U.S. labor force (Bureau of Labor Statistics, 2017). By 2060, approximately 20% of the total national population is expected to be foreign-born (Colby & Ortman, 2015). Some raise concerns about the potential loss of jobs by American-born workers to immigrants (Hoban, 2017). Others argue the rising immigrant population will increase opportunities for U.S.-born workers, as immigrants frequently perform low-skilled jobs that are complementary to, and increase the productivity of, work performed by other Americans (Greenstone & Looney, 2012). However, many immigrants also hold advanced degrees, particularly in STEM fields (Solis, 2011), and could therefore play a crucial role in meeting the demand for highly skilled workers.

As working Baby Boomers draw closer to retirement age, there is concern over the loss of the knowledge and skills of the overall labor pool (Burke & Ng, 2006). Globally, the ratio of non-working age people to working age people appears to be on the rise (Bakhshi, Downing, Osborne, & Schneider, 2017). This trend may be counterbalanced by policy changes that raise...
retirement ages or provide incentives for older workers to remain on the job (Lerman & Schmidt, 1999).

The overall aging of the population has implications for available jobs. Jobs in healthcare and the production of goods and services targeting the needs of older citizens are on the rise (Singh, 2015). However, the influx of highly-educated Millennials into the workforce has its own implications. These workers are anticipated to bring a new set of expectations of their employers, including demands for improved working conditions and human resources policies (National Academies of Sciences, Engineering, and Medicine, 2017). At the same time, there is rising concern regarding this full subpopulation’s preparedness with the skills required in the ever-changing world of work (KRC Research, 2014).

**Education and Training**

It has been estimated that most children entering primary schools today will work in job types and roles that don’t yet exist and that will be characterized by the need for not only technological, but also social and analytical skills (World Economic Forum, 2016). It is anticipated workers of the future will hold an increasing number of jobs over their lifetime (Pompa, 2015). These factors, coupled with increasingly rapid technological change, will necessitate a continuous process of education and training throughout these future workers’ careers (Karoly & Panis, 2004). This suggests the need for consideration of both the education and training offered to students prior to their entry into the paid labor force, as well as how systems for continued education and training will be implemented and sustained.

Numerous innovative approaches to preparing students with the in-demand middle level skills needed to perform jobs of the future are expanding in their implementation. Career and technical education (CTE) programs, apprenticeships, early college high schools, and career academies are among the approaches that seek to bolster the skill levels of Americans entering the workforce for the first time (Joint Economic Committee Democrats, 2018). On-the-job training (OJT) models are another innovative approach that provides incentives to employers to hire lower-skilled workers and offer them targeted training while they engage in paid labor, as well as offer continued training to allow for career advancement (Kobes, 2013).

**Equity Issues**

Many anticipate the trends of globalization and automation will lead to increasing inequality, as wages for highly skilled workers rise while low- and unskilled workers will compete with both automation and workers located in other countries (The Foundation for Young Australians, 2017). Other areas of concern regarding equity in the workplace relate to gender and age. Women who seek to both parent and work continue to face potential wage reductions and loss of skill development when they take time off for family leave (O’Marah, 2018). Some anticipate women will be disproportionately impacted by job losses due to automation (Hayasaki, 2017). Aging workers may face threats to their continued employment over issues related to healthcare costs and age-related disabilities (National Bureau of Economic Research, 2018).

**New Social-Oriented Jobs**

Futurists envision new positions will be created to do work that has never been done before. Most of these jobs will develop in response to shifts in the marketplace or they will be created because of advancing technologies. Envisioned jobs include those that harness the power of social media to create tailored experiences for customers or clients (Wagner, 2010). With more companies using social media to connect with customers and to expand their presence in the market, employees’ experience with and understanding of social media will be increasingly valued by employers (Kumar, Bezawada, Rishika, Janakiraman, & Kannan, 2016). Companies
and organizations will need to monitor, maintain, and improve their online presence, and new positions will likely be created for that purpose (University of Kent, 2018).

**Workplace of the Future**

When examining the workplace of the future, Frey and Osborne (2013) convened human experts in machine learning to classify a subset of jobs according to the likelihood of their “automatability.” Through analysis of O*NET variables as proxies for three irreplaceable attributes (i.e., perception and manipulation, creative intelligence, and social intelligence) they developed a model to predict the automatability of the full set of O*NET occupations. Results indicated that 47 percent of U.S. employment can be classified as high risk for automation within the next decade or so.

The Guardian’s Workplace Benefits Study (2017) defines four top trends impacting the workforce in 2018 and beyond. Each of these trends is related to technology:

- Technology is enabling an on-demand workforce;
- Automation is requiring an enhancement of workforce skillsets;
- Employers are reinventing talent recruitment; and
- Varying workplace demographics require different strategies for adoption.

**Workplace Culture**

**Agile Workforce**

As organizations are required to respond quickly to changes in an increasingly globalized and technologically advanced world, they seek an agile workforce that is similarly capable of responding to unanticipated change with speed and flexibility (Breu, Hemingway, Strathern & Bridger, 2001). Workers of the future may be expected to rotate among a variety of roles and tasks, as employers seek to find the skills needed for a specific task at a particular time (Wadors, 2018). As companies leverage a variety of work models (e.g., ad hoc teams, crowdsourcing, independent contractors) to meet their needs, workers may find themselves entering into many different types of nontraditional work arrangements (Green, 2014). Andrew, Ip, and Worthington (2014) expect an increase in distributed work places. Continual reskilling will be a key element in sustaining high levels of agility (Lyons, Blitz, & Whittall, 2017).

**Less Structure and Predictability**

Careers have been traditionally viewed as a progression of jobs, often upward through a predictable, hierarchical structure (Lyons, Schweitzer, & Ng, 2014). Careers of the future will likely unfold in less hierarchically structured environments, where there will be increased interconnectedness among departments and where individuals may assume different job roles depending on the context of the work at hand (Heerwagen, 2016). Job tasks themselves are expected to be less structured and predictable as new technologies replace once rote and predictable duties with ones that require abstract thinking and flexibility (National Academies of Sciences, Engineering, and Medicine, 2017).

**Sharing Economy**

More and more, modern day consumers and workers engage in short-term economic transactions around services that involve sharing some material good (e.g., car, living space) or skill for monetary compensation (Sundararajan, 2016). Also referred to as the gig economy, platform economy, access economy, or collaborative consumption, this sharing economy is
anticipated to increase exponentially over the coming decades (Yaraghi & Ravi, 2016). Such work arrangements have both potential positive and negative consequences for workers of the future. It can be argued that individuals will be empowered by the sharing economy to go into business for themselves and gain returns on their assets. On the other hand, the sharing economy removes protections that workers have enjoyed under more traditional work arrangements (Lamberton & Rose, 2012).

**Continuous Learning**

McKinsey & Company (2017) recommend that workers of the future be prepared to be lifelong learners. McKinsey Global Institute (MGI) partner Susan Lund explained, “For young people today, what’s clear is that they’re going to need to continue to learn throughout their lifetime. The idea that you get an education when you’re young and then you stop and you go and work for 40 or 50 years with that educational training and that’s it—that’s over. All of us are going to have to continue to adapt, get new skills, and possibly go back for different types of training and credentials. What's very clear is that what our kids need to do is learn how to learn and become very flexible and adaptable.”

Guardian (2017) recommends that employers address the need for continuous learning through experiential, retraining, and cross-training programs, as well as mentoring, e-learning opportunities, and tuition assistance.

Ross (2016) opines the U.S. adoption of free education until the age of 18 was appropriate as long as a high school graduate could get a job in a “port, factory, mine or mill—a middle class job.” However, in the information age, he suggests we know the pace of change demands that we be lifelong learners.

**Flexible and Non-Traditional Career Paths**

Predictions regarding future career paths are wide ranging. Popular “wisdom” has long asserted that younger generations no longer expect to join an employer after high school or college and stay with that same employer until retirement. Lyons, Schweitzer & Ng (2015) analyzed the career mobility patterns of four generations and found that job mobility increased with each successive generation. Specifically, “The magnitude of the differences was large, as Millennials [born 1980 or later] had almost twice as many job and organizational moves per year as the generation Xers [1965-1979], almost three times as many as the Boomers [1946-1964], and 4.5 times as many as the Matures [born prior to 1946]” (page 16). However, this change in job mobility does not reflect an increase in employee turnover from one employer to another, but rather increased movement through various positions within a company. They postulate that technology, among other factors, may make some positions obsolete. The authors conclude the traditional career model is still strong and the “oft-cited truisms about the ‘new’ or ‘modern’ careers may be exaggerated” (page 18).

Intuit & Emergent Research (2017) predict that by 2021, 9.2 million American workers will derive at least some of their income as independent contractors operating within a “gig economy”—situations in which organizations establish short-term contracts on an as-needed basis. This is a substantial growth projection relative to the 3.9 million in 2016. McKinsey Global Institute (2016) estimates that 20–30 percent of individuals of working age in the U.S. and the European Union conduct independent work.

This trend is facilitated by technology that allows a job incumbent to be geographically distant from the employer; the advantages to an organization of selecting the best candidates for a given project, without a long-term commitment; and the ability to increase and decrease staff
levels as demand warrants. This is further enabled by current and planned features in job-employee matching software such as Monster.com, Aftercollege.com, and Taskrabbit and networking sites such as LinkedIn (Brynjolfsson & McAfee, 2016).

**Interdisciplinary Teams**

Based on research by Burkus (2016), some organizations encourage employees to engage in more face-to-face communication in an effort to increase problem solving and decision making efficiency (as cited in Colbert, Yee, & George, 2016). As a result, workplaces evolve to provide more flexible space for collaborating and working in teams (Giang, 2015). Experts from Unum Limited's Futures100 network (2014) foresee more conversation and debate, either face-to-face or on conversation-based platforms. Employees will need to blend skills and disciplines when working with others. They will collaborate with each other rather than compete. Workers will need listening skills and to display empathy, and build relationships to enable collaborative and interdisciplinary ventures.

**Summary of Themes of Work and Workplace of the Future**

When it comes to work of the future, change is the only certainty. However, this review of relevant literature points to some overarching themes that provide a solid base for making predictions about the world of work that today's kindergartners will need in 2030 when they graduate from high school. This world will likely look very different from the world of work their parents were prepared for, both in terms of the available jobs and the work environment in which those jobs are carried out.

Jobs of the future will undoubtedly involve technology. From searching job openings, to performing job tasks, to receiving professional development, interacting with new and emerging technologies will be a distinctive feature of future jobs. Fields that had previously been quite separate may be blended in new ways, and existing jobs may be blended with new technologies to create positions we've never seen (think: space junk recyclers!).

The high school graduates of 2030 will set out on a career pathway characterized by change. Whether they work independently through the gig economy, or move among multiple employers or across multiple departments or projects, workers of the future will likely find themselves part of an increasingly diverse and dispersed workforce. Jobs will be continually evolving to meet changing demands and to incorporate the latest innovations. Ongoing training will be a necessary component of future jobs. Employees will need to adapt and embrace life-long learning to be successful in the workplace.

With some sense of what the future holds for work and the workplace, it becomes clear expected changes in jobs and job environments will correspond to changes in associated skills. A key next step to ensuring that students graduate high school in 2030 prepared for the next step on their postsecondary pathway is to identify the skills and abilities needed to successfully perform the jobs of the future.
References


Skills of the Future – 2030

Overview of Skills of the Future

Students completing kindergarten in 2018 will graduate from high school in 2030. As the work and workplace of the future change, so will some of the skills students need for success in postsecondary activities. In *Work of the Future – 2030*, Gribben, Becker, and Dickinson (2018) described research related to potential changes in the workplace and forecasts of jobs of the future. Although we do not know with certainty what types of jobs will be available in the future, researchers and business analysts use trends to predict the types of jobs and skills that they expect high school graduates will need for employment in the future (e.g., Grover, 2018; Lara, 2018; McKinsey & Company, 2017; P21, 2016).

This review follows that of the *Work of the Future – 2030* and answers a range of questions about the landscape of postsecondary skills in 2030. What skills will students need following graduation from high school in 2030? Are there skills common across jobs? What skills do secondary students need either for matriculation into college or entering the workforce?

Skills high school graduates in 2030 will need to succeed along postsecondary pathways must correspond to the jobs of the future. Technology is expected to play a large role in future jobs. With the advent of the digital age, there is a recent emphasis on skills required to use and interact with new devices and applications. Employers will need programmers and innovators to develop new technologies to tackle more difficult challenges and improve efficiency and cost effectiveness. High school graduates of 2030 will likely find themselves part of an increasingly diverse and dispersed workforce. As the workplace and postsecondary institutions become more distributed and global, employees and students will need facility with collaboration tools as well as socio-emotional skills for working with diverse colleagues. Jobs will continually evolve to meet changing demands and to incorporate innovations. Employees will need to adapt and embrace life-long learning. Automation will replace some jobs and disrupt certain industries; future jobs are likely to require complex problem solving and troubleshooting which are not easily automated. Skills that enable individuals to work with and use technology, communicate with others, and continually adapt and learn will be necessary for the high school class of 2030.

In this literature review of the skills of the future, we provide a brief introduction to college and career preparedness. We present a structure for organizing the discussion of skills of the future followed by research on each of the skills. Similar to the *Work of the Future – 2030* literature review, we include projections of shifts in future skills. We discuss skills expected to increase in demand through 2030 and new skills expected to be added. The report concludes with a summary of themes of skills for the future.

Integrated Framework for Postsecondary Preparedness

There has been much discussion and work looking at college and career preparedness either by bifurcating college and career preparedness or by assuming them to be the same (ACTE, 2010; Conley, 2011; National Center for O*NET Development, 2018; P21, 2016). In addition, global statements frequently are made about preparedness without considering context. Patelis (2018) has proposed an integration of popular career and college preparedness frameworks to address the limitations in other organizing schemes to cover both career and college.

Focusing on the skills at the intersection of career and college preparedness, we use a unified framework to represent postsecondary preparedness. This offers a way of thinking about the overlapping nature of the skills needed for career and college in a more integrated manner.
rather than as separate or redundant constructs. The National Research Council (NRC, 2011; Pelligrino & Hilton, 2012) organized postsecondary skills into three categories: cognitive, intrapersonal, and interpersonal. These clusters encompass knowledge and skills\textsuperscript{41} needed for life after high school. To address the more complex and multidimensional skills of the future, however, we have added another category – blended skill sets.

As the NRC defined them, cognitive skills for the 21st century involve (a) cognitive processes and strategies, (b) knowledge, and (c) creativity. Cognitive processes and strategies encompass critical thinking, complex problem solving, and analysis and interpretation. Knowledge covers academic areas such as reading, writing, and science, technology, engineering, and mathematics (STEM). With the ubiquity of digital devices, skill in using or developing digital tools is an important addition to the set of cognitive skills of the future. Creativity includes innovation and creative skill sets.

Interpersonal skills, sometimes called social skills, require complex communication, teamwork, and collaboration. These range from communication and collaborative problem solving to cooperation and perspective taking. With expected increases in diverse and globally dispersed workplaces of the future, we have added cultural awareness and sensitivity to the interpersonal skills needed in the future.

Intrapersonal skills cover intellectual openness, and work ethic and conscientiousness. Intellectual openness includes adaptability, personal responsibility, and continuous learning. Work ethic and conscientiousness includes initiative, productivity, and professionalism.

The following sections look at the future of (a) cognitive skills, (b) interpersonal skills, (c) intrapersonal skills, and (d) blended skill sets. Table 1 presents a list of the skills within each of the categories. The skills discussed here are not intended to be comprehensive. Rather, we have focused the discussion on skills that figure most prominently in recent thinking of the future of work.

\textsuperscript{41} Knowledge refers to what a person knows and understands. Skill refers to what a person can do.
Table 1. Postsecondary Preparedness Skills for 2030

<table>
<thead>
<tr>
<th>Cognitive Skills</th>
<th>Interpersonal Skills</th>
<th>Intrapersonal Skills</th>
<th>Blended Skill Sets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foundational Academics</td>
<td>Communication (including Listening, Conversation, and Persuasion)</td>
<td>Time Management</td>
<td>Learning Agility</td>
</tr>
<tr>
<td>STEM</td>
<td>Relationship Building</td>
<td>Efficiency</td>
<td>New Media</td>
</tr>
<tr>
<td>Critical Thinking</td>
<td>Cultural Sensitivity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complex Problem Solving</td>
<td>Understanding Other People’s Perspectives</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Creativity</td>
<td>Collaborative Problem Solving</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Innovation</td>
<td>Social and Emotional Intelligence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digital Tools</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistical Literacy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computational Thinking</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Cognitive Skills**

- **Foundational Academics**
  - Basic skills in education, namely literacy and numeracy, are well defined, taught, and measured. They are foundational to the acquisition of knowledge and skills (Peterson et al., 2001) and the performance of most tasks (e.g., Durak & Saritepeci, 2018). Economic research shows that improving basic skill proficiency has a dramatic effect on important societal outcomes, such as wage growth (McIntosh & Vignoles, 2001) and social development (Hanushek & Woessmann, 2008), further demonstrating the criticality of these skills to important outcomes.

  As work evolves over the coming decades, literacy and numeracy will remain important, although the way they are used may change. For example, tasks requiring basic literacy and numeracy, such as data entry and cashiering, will likely become automated, thereby reducing the demand for these skills by as much as 25% from 2016 to 2030 (Bughin et al., 2018). As work becomes less structured and predictable and more team- and project-based, literacy will remain critical for effective acquisition and communication of information. For example, the demand for more advanced literacy and numeracy skills, such as are used in analytic and communications activities, is expected to increase by almost 10% (Bughin et al., 2018).

  Other academic areas beyond literacy and numeracy include mathematics and science. These have become increasingly important and are often grouped together with technology and engineering to form their own category of Science, Technology, Engineering, and Mathematics (STEM).

- **STEM**
  - Skills in the STEM areas are frequently included in discussions of the future of work. While great effort is currently being put into projecting and developing STEM skills for the future, the concept of STEM as an educational discipline dates to at least two National Science Foundation reports...
in the 1990s (Advisory Committee to the National Science Foundation Directorate for Education and Human Resources, 1996; 1998). Since that time, researchers worldwide have attempted to define STEM skills and strategies for developing them in the current, and future, workforce.

While many of these discussions project shortages of STEM skills, other research suggests that there is both a shortage and a surplus of STEM skills, driven by occupational and geographic mismatches between individuals’ skills and specific job openings. STEM is a very broad concept and not all skills within it are equal in demand. Xue and Larson (2015) noted that, due to the heterogeneous nature of STEM, some skills are in short supply. These include very rare skill sets (e.g., Ph.D.-level nuclear engineering), as well as more common (e.g., data science, software development) and overlooked STEM skills (e.g., skilled trades). Xue and Larson also identified surpluses of STEM skills, such as biomedical and chemical engineering, as well as geographical variation in supply and demand (e.g., software engineering skills are in high demand in California).

Other researchers point out that basic STEM literacy will be critical even for non-STEM jobs. In fact, some definitions of basic skills include STEM literacy as a basic skill or assume that most job candidates will have at least basic proficiency with STEM skills (Cunningham & Villaseñor, 2016; Roberts & Bybee, 2014; Zeidler, 2014) due to its perceived relevance for a world of technology-immersive work.

There is disagreement regarding the profile of skills defining STEM. Some definitions delineate specific STEM disciplines along with associated higher-order cognitive skills. Carnevale, Smith, and Melton (2011) and Jang (2015) empirically identified the skills associated with STEM occupations using the Occupational Information Network (O*NET42) taxonomy. O*NET classifies its nearly 1,000 occupations into 5 job zones describing how much preparation (education) is required, job climate (e.g., working conditions, recognition, independence), and potential for future growth. Results of the STEM occupation crosswalk included a wide range of skills including both content knowledge (e.g., math, chemistry, and other scientific and engineering fields) and a broader set of cognitive skills relevant to the future of work (e.g., complex problem solving, deductive and inductive reasoning, mathematical reasoning, and facility with numbers).

Other definitions focus on the interdisciplinary components of STEM as they apply across disciplines and occupations:

STEM skills and knowledge are interdisciplinary in nature, being based on the integration of the formerly discrete disciplines of science, mathematics, engineering and technology. The aim of STEM skills is to enhance people’s competency in work and/or life and more generally respond to societal demands on technology.

STEM skills belong to the group of technical skills. They are a combination of the ability to produce scientific knowledge, supported by mathematical skills, in order to design and build (engineer) technological and scientific products or services. Although STEM skills overlap with basic and higher order cognitive skills, they merit separate treatment in a policy-oriented context in order to target specific requirements in the education and labor market. STEM skills and knowledge

---

42 O*NET (https://www.onetonline.org/) is developed under the sponsorship of the U.S. Department of Labor/Employment and Training Administration (USDOL/ETA).
cannot be directly measured by current discipline-specific classifications.
(Siekmann & Korbel, 2016)

Siekmann’s (2016) “House of STEM” model integrates these perspectives. The “House of STEM” includes basic numeracy, literacy, and socioemotional (e.g., curiosity) skills as a foundation for the kinds of technical and higher-order cognitive skills identified by Carnevale et al. (2011).

In addition, other researchers advocate for broader conceptualizations of STEM that include skills from the arts and humanities (i.e., STEAM). These models reflect the importance of visual and verbal communication to both the discipline-specific technical skills and the higher-order cognitive skills comprising the House of STEM (Land, 2013). For example, modern approaches to data analytics rely heavily on data visualization, a skill that clearly draws upon the arts. The arts also help develop higher-order cognitive skills, such as critical thinking, creativity, perspective taking, and divergent thinking, that are critical to scientific advancement and innovation (Daugherty, 2013; Kahn & Zeidler, 2016). These researchers caution that focusing solely on a narrow definition of STEM skills may not adequately prepare us for the future.

**Critical Thinking**

Critical thinking is important for learning and for evaluating new information (Mumford, Peterson, & Childs, 1999), making it vital to performing work that does not follow prescribed rules (i.e., work that is not likely to be automated). In a future world of work characterized by expanding technology and artificial intelligence, humans’ unique capacity for critical thinking will set them apart from machine learning algorithms (Barnett, Lawless, Kim, & Vista, 2017; Pistrui, 2018). In fact, job recruiters across the major industries have identified critical thinking skills among the less common, but more desired skills among business school graduates (Levy & Cannon, 2016).

Glaser (1985) viewed critical thinking as a process involving careful consideration of a problem through the application of logical inquiry methods. McPeck (2016) identified the critical features of critical thinking, six of which distinguish critical thinking from other higher cognitive skills:

1. Involves the evaluation of statements/evidence and the methods used to derive them
2. Is limited in application to the context of a specific discipline, field, or problem; it cannot be learned or applied in the abstract
3. Does not require the thinker to reject established norms or ideas
4. Is informed by evidence but requires judgment
5. Includes active problem solving, not just evaluation
6. Is not redundant with logic or rationality.

**Complex Problem Solving**

In recent years, employers have consistently rated problem solving as the most essential competency for career preparedness (National Association of Colleges and Employers, 2017; Thompson, 2016). CEOs from across industries and around the globe expect work environments to continue to grow in terms of complexity (IBM, 2010), and complex problem solving is anticipated to replace physical abilities and core skills in a wide variety of jobs of the future (World Economic Forum, 2016). Thirty-six percent of all jobs are expected to require complex problem-solving as a core skill by 2020 (Thompson, 2016). The ability to address
complex problems is particularly important in work environments characterized by rapid change (Middleton, 2002).

Complex problems are those that involve multiple goals along with many possible courses of action, all of which may shift in an environment that is dynamic (Fischer, Greiff, & Funke, 2012). Complex problem solving includes not only cognitive, but also emotional and motivational elements. It is a dynamic process in which the pathway to the solution might be more informative about the problem solver than reaching the solution itself (Dorner & Funke, 2017).

Creativity

U.S. CEOs are looking for employees with creativity (Ryan, Sapin, Rao, & Ampil, 2018). Creativity is a skill often identified as important in a rapidly changing workplace, and a capability that machines do not possess. But the assumption that creativity is a purely human skill has been challenged. Early creativity assessments measured divergent thinking or resistance to functional fixedness. These assessments included tasks such as presenting an object and asking the examinee to generate as many uses of that object as possible, within a specified timeframe. Individuals who generated a larger number of unique uses, regardless of elegance or complexity, received higher scores. This task measured the ability to generate novel, or heretofore nonexistent, ideas.

One technique for creating novel ideas is to simply produce combinations of familiar ideas. For example, a common brainstorming exercise to generate ideas for creative writing is to create index cards—either electronic or physical—each with a noun or adjective, and then pick two or three cards at random to create an unusual combination to spark an original idea. Computers are certainly well suited to this sort of rapid, rote activity. However, while this approach demonstrates an ability to generate novel ideas it may result in nonsensical combinations.

Boden (2003) builds upon this definition of creativity as “the ability to come up with ideas or artifacts that are novel and valuable” (emphasis added). Similarly, Mumford (2003) describes creativity as “new and useful.” Under this framework, creativity requires the generation of unique ideas as well as the ability to evaluate those ideas. Evaluating creative ideas is a complex and domain-specific task that requires a wealth of information. Not only must this conceptual space include sufficient details about the field to evaluate whether ideas are sensible, but metrics are necessary to assess the value of an idea. These values are specific to content domains and may change over time.

Frey and Osborne (2013) point out that research literature reveals examples of software generating and evaluating creative solutions. For example, Harold Cohen produced a drawing program, AARON, in the mid-1970s in an attempt to answer the question, “What are the minimum conditions under which a set of marks functions as an image?” Over decades, AARON evolved into an extensive producer of visual art with an internal feedback system that has been displayed in art galleries (Cohen, 1995).

Similarly, David Cope produced the initial version of the Experiments in Music Intelligences (EMI) software program in the early 1980s to analyze the style of a given musical composer (e.g., Bach, Mozart, Prokofiev) and then generate an original composition in the same style (da Silva, 2003). This software required domain-specific knowledge of tone systems, phrase structure and length, rhythm, movement structure, etc. and the ability to evaluate the balance between unity and variety.
Despite these domain-specific examples of computer generated creative products, Frey and Osborne (2013) conclude that “it seems unlikely that occupations requiring a high degree of creative intelligence will be automated in the next decade” (p. 26).

Brynjolfsson and McAfee (2011) suggest that the combination of current economic trends, specifically job growth trends and wage stagnation at the median range, provides opportunities for creative entrepreneurs. While increased technology may eliminate some career paths, it also opens potential paths to create ways to use technology to support mid-skilled workers and add value.

**Innovation**

The Future Laboratory conducted a study to identify trends in workplaces, employers, and employees (Unum United, 2014). They interviewed a range of experts from its Futures100 network—including academics, authors, scientists, and social scientists—and also surveyed 1,000 employees for their perspectives on the trends identified via these interviews. Respondents indicated that future workplaces will be people-centric and will foster innovation. Employees will need to blend skills and disciplines.

Innovative employees can only thrive in an organization that accepts innovative ideas. While innovation has been touted for years as an important trend, recent studies call specifically for “responsible” or “disciplined” innovation. Sull (2015) appeals for the implementation of a simple set of rules to serve as constraints in the innovation process. He acknowledges that despite guardrails designed to manage innovation, some failures are still inevitable, but that incorporating discipline serves to increase efficiency and improve the odds of successful innovations. He characterizes an appropriate set of simple rules as (a) few in number, (b) applicable to a well-defined activity or decision (rather than a broad corporate principle that is too vague to be actionable), (c) tailored to the culture and norms of the organization, and (d) sufficiently flexible to allow creativity and discretion. He suggests that an organization might employ rules for various purposes, including to select innovations, define how to innovate, and help community members innovate together. He cites successful innovators ranging from Zumba Fitness to the Defense Advanced Research Projects Agency (DARPA).

A corporate reputation for innovative thinking also can attract new employees. Brown and Martin (2015) describe the success of Innova Schools, which brought affordable education to Peru. Innova drew upon the results of a stakeholder study to develop a technology-enabled education system that valued the “guide on the side” rather than the traditional “sage on stage” model of instruction. Invited, ongoing feedback from school leaders, teachers, and parents helped to introduce continuous improvements, some of which were fundamental shifts from the expected outcomes. This approach resulted in scalability and job growth. The authors note that “Because Innova had a reputation for innovation, teachers wanted to work there, even though it paid less than the public [school] system” (p. 9).

**Digital Tools**

While the construct of skill with digital tools is not well defined, there are two aspects of digital skills that are unique: the skills required to (a) create and (b) use digital tools. The creation of digital tools, such as artificial intelligence and machine learning, requires STEM, analytic, and computational thinking skills. More specific digital skills, such as web development, are likely to rapidly evolve with the advent of new technologies, methods, and languages.

Basic digital skills are useful for developing other skills. For example, some medical training now relies on virtual reality simulations (e.g., Azarnoush et al., 2015), including to assess trainee
Virtual reality is also being used therapeutically, such as to develop emotional skills in children with autism spectrum disorders (Lorenzo, Lledó, Pomares, & Roig, 2016).

**Statistical Literacy**
Many expected changes in the nature of work relate to enhanced access to, and use of, data to make decisions. While some basic statistical processes can be automated with algorithms and machine learning, Gal (2002) argued that citizens must have a basic understanding of the concepts underlying statistical reasoning and terminology to interpret data, evaluate sources, and make decisions, including:

1. Basic statistical and mathematical methods and terminology
2. Foundational statistical concepts, such as probability and variability
3. Inferential reasoning
4. World knowledge, to aid interpretation and evaluation of findings.

Several trends will lead to increased demand for employees in all occupations who can (a) effectively use data, (b) understand how to visualize and manipulate data, and (c) draw conclusions from data. First, data are becoming more accessible. Brynjolfsson and McAfee (2011) pointed out that “information doesn’t get used up even when it’s consumed…. and once a … body of information is digitized …. it can be copied infinitely and perfectly, and distributed around the world instantly and at no additional cost. This is nothing like the economics of traditional goods and services” (p. 73). The ubiquity of data and the relative ease with which it can be distributed and shared opens the possibility of extensive data analytics. Bonney et al. (2009) discussed the role of such “citizen scientists” in the advancement of both educational and scientific outcomes.

Second, access to new technology has further increased the importance of statistical literacy. Big data, artificial intelligence, and machine learning play an increasing role in daily life. While these technologies will replace some existing human tasks, they are expected to create new skill requirements, such as the ability to train algorithms (e.g., generating models for machine learning used in automated item scoring or image recognition used for automated driving), explain how they work, and keep them operating (Wilson, Daugherty, & Morini-Bianzino, 2017). Statistical reasoning skills will underlie all of these roles.

Finally, statistical literacy can also require specific skills, such as data visualization (Fox & Hendler, 2011). Hampton and colleagues (2017) created a taxonomy of skills for data-intensive research that includes five skill areas (data management and processing, software skills, analysis skills, data visualization, and communication and collaboration for results dissemination). While their focus was on environmental science, this taxonomy applies equally to work in other research-oriented domains as well.

**Computational Thinking**
Computational thinking can be thought of as a special case of analytical thinking, one that draws specifically on the ability of computers to abstract and automate problem solving (Wing, 2008). It is now pervasive in most analytical disciplines (Beheshti et al., 2017).

Hu (2011) defined computational thinking as:
thinking to solve problems, automate systems, or transform data by constructing models and representations, concrete or abstract, to represent or to model the inner-working mechanism of what is being modeled or represented as an information process to be executed with appropriate computing agents. Such thinking is necessarily:

- logical, to capture what is essential to the models or representations;
- algorithmic, to step-wise define or refine operational processes;
- scientific, to gain understanding of models’ capabilities, learn how to use them with maximum efficiency, and explore the effects of the computation in the original problem domain.
- mathematical, to be able to show the correctness of algorithms, specify precisely the functionality of a software system, measure the quality of what we do in a process of computation, and deal effectively with the complexity of the models and representations by exploring more effective and efficient alternatives;
- analytical, to model with purpose, assumptions and viewpoints, evaluate and adjust the models and representations by prototyping, and study their implications and consequences;
- engineering-oriented, to design the models and representations against known constraints and practical concerns, and to plan, execute, manage, and evaluate the process of computation in order to improve our capability and maturity level; and
- creative, to model the unthinkable.

It is important to note that computational thinking is not synonymous with coding or programming skills, but encompasses understanding of computational concepts, practices, and perspectives (Lye & Koh, 2014). Psycharis (2018) argued that computational thinking integrates mathematics, computer science and knowledge in one or more subject areas to solve complex problems. In measuring criterion validity of the Computational Thinking Test, Román-González, Pérez-González, and Jiménez-Fernández (2017) found statistically significant correlations between computational thinking and spatial, reasoning, and problem-solving abilities.

**Interpersonal Skills**

**Communication**
Communication is, at its core, an exchange of information, whether linguistic or non-linguistic, and is widely considered a key competency in both postsecondary education and workplace contexts (Brink & Costigan, 2015). Effective communication, defined as the ability to synthesize and transmit ideas, is among the critical skills needed by employees at all levels of organizations (American Management Association, 2012). Communication in the future world of work will require the ability to work with emerging technologies, along with the more traditional elements of communication such as listening, initiating and engaging in conversation, and persuading others.

**Listening**
Listening has been identified as the most important oral communication skill for successful job performance across a range of workforce samples (Brink & Costigan, 2015). Listening can be categorized into four major types:
1. Active- giving full attention when others are speaking
2. Involved- giving most of one’s attention to the speaker’s words and intents
3. Passive- Receiving information rather than being an equal partner in an exchange
4. Detached- Withdrawn from the speaking-listening exchange such that one is the object of the message rather than the receiver (Pearce, Johnson, & Barker, 2003).

Active listening, in particular, is sought after by potential employers as it helps to create a positive work culture and supports collaboration, which in turn spurs innovation (Nowogrodski, 2015). An active listener fully concentrates on what is being communicated and provides both verbal and nonverbal feedback in response. Examples of verbal feedback include providing positive reinforcement, remembering prior details, asking relevant questions, paraphrasing what the speaker has said, and requesting clarification. Examples of nonverbal feedback during active listening include smiling, making eye contact, maintaining posture, mirroring facial expression, and maintaining focus (SkillsYouNeed, 2018).

Conversation
Conversation among team members, whether virtual or face-to-face, is anticipated to be an expanding feature of future jobs (Gribben, Becker, & Dickinson, 2018). Conversation skills are important because they contribute to an organization’s shared understandings, which may be critical for the agile decision-making that is characteristic of the workforce of 2030 (Heidema, 2017). Conversing skill has been rated among the most important oral communication skills, typically rated between listening and presentation in terms of importance for job success (Brink & Costigan, 2015). Employers seek employees with conversation skills because they will contribute positively to the workplace culture by promoting dignity and increasing motivation (Macaulay, 2014).

Conversation goes hand-in-hand with listening, such as knowing when it’s time to listen and when it’s time to talk, and gauging one’s delivery based on mindful listening to the other person’s message (Macaulay, 2014). Conversation skills also include staying organized, and being strategic about the information one both conveys and takes away from the interaction (Coplin, 2003). In the digital age, workers are increasingly engaging in multiple communications, often simultaneously, and therefore run the risk of tuning out important conversations as they seek to filter all of the information received. The ability to engage in authentic conversations on social media platforms therefore becomes a skill of its own (Lombardi, 2014).

Persuasion
Persuasion is a uniquely human skill, one that is expected to withstand the wave of increased automation (Luckin, Baines, Cukurova, Holmes, & Mann, 2017) and one that is increasingly in demand (Deloitte, 2016). Also, as the future world of work will be characterized by increasing diversity and geographic dispersion, the ability to persuade people from a variety of backgrounds will be valued in the workplace (Martin, 2010). Persuasion is an important skill for any job role that involves managing customer or client relationships, or managing other employees (Dellaert & Davydov, 2017).

Persuasion skills include (a) making an assessment of the individual or group one is trying to persuade, (b) establishing rapport with them, (c) communicating the benefits of
the proposed course of action, (d) actively listening to any counterarguments, (e) clearly presenting counterpoints to these arguments, (f) recognizing any limitations of the original course of action, (g) modifying the course of action as needed, (h) reaching terms with the person being persuaded, and (i) following up to ensure that they are still on board with the agreed upon course of action (Doyle, 2018).

**Relationship Building**

As the workplace of the future relies more and more on teams, both virtual and in-person, relationship building skills become more valuable. Building relationships has benefits for employees, teams, leaders, and organizations, such as building trust, boosting morale, and improving decision making (Pauleen, 2004). The ability to forge positive relationships in the workplace is key for an individual’s job satisfaction, and is an essential building block in the creation of a collaborative work environment.

Relationship building is characterized by listening to others and encouraging them to share their thoughts and feelings (Lievens & Sackett, 2012). Communication skills therefore play a major role in relationship building, but also key are things such as following through on commitments and being considerate of others’ feelings and perspectives (Tingum, 2018). Relationship building skills are characterized by willingness to share one’s knowledge and expertise; providing quality feedback to others; supporting others’ work while also bringing in others to help with their own work; and engaging in ongoing, friendly interactions inside and outside the workplace (Garfinkle, 2018).

**Cultural Sensitivity**

Cross-cultural competency will be a core skill in most organizations of the future, as employees will need to be able to identify shared values to work effectively with increasingly diverse coworkers (Davies, Fidler, & Gorbis, 2011). In the context of the workplace, cultural sensitivity includes the ability to work effectively alongside someone from a different cultural background who may approach workplace behaviors differently (Sherman, 2018). Culture-based misinterpretations can have implications for the success of collaborative efforts (Blanding, 2012).

Coworkers from different cultural backgrounds may engage in different behaviors and hold different work-related values. For example, employees from individual-oriented cultures may approach work tasks differently than someone from a group-oriented culture (Heggertveit-Aoudia, 2012), which may influence behaviors such as how employees participate in meetings, the amount of time they spend socializing, and whether they provide feedback or otherwise publicly express opinions (Knight, 2015). Increased cultural sensitivity could help mitigate such differences.

At its most basic level, cultural sensitivity requires knowledge and understanding of other cultures (Lutz, 2017). Cultural sensitivity may also involve taking an interest in another culture, recognizing cultural differences, and then changing one’s own behavior to show respect for the other culture (Hammer, Bennett, and Wiseman, 2003). Recognition of one’s own biases is also an element of cultural sensitivity (Loue, Wilson-Delfosse, & Limbach, 2015).

**Understanding Other People’s Perspectives**

Increasing levels of collaboration among diverse teams in the workplace will boost the value of perspective taking as a job skill. Perspective-taking refers to the ability to take on another person’s point of view. It is an active and goal-directed process that involves trying to
understand the thoughts, and feelings of another, as well as the motivations behind them (Parker, Atkins, & Axtell, 2008).

Situational awareness and personal awareness are two key components of perspective-taking. Situational awareness refers to understanding the context in which another person is acting. Personal awareness refers to understanding what the other person brings into that context (Goulston & Ullmen, 2013). Other building blocks of perspective-taking include being aware of others, regulating one’s emotions and empathy, being able to successfully “read” other people, and correctly interpreting what others are trying to communicate (Campbell, 2016).

Collaborative Problem Solving
The demand for collaborative problem-solving skills is anticipated to experience high levels of growth in the future. It is defined as the ability to engage effectively with two or more people to solve a problem through shared understanding and effort, and pooled knowledge and skills (Luckin, Baines, Cukurova, Holmes, & Mann, 2017). Collaboration will be key as increasingly complex problems will not be solved by one specific field of expertise, but rather will require working with others from different disciplines (Davies, Fidler, & Gorbis, 2011). Collaborative problem solving was recently added to the skills measured by the Program for International Student Assessment (PISA), a reflection of its significance as a desired skill. As the workplace of the future will be characterized by increasing amounts of teamwork, being able to collaborate to solve problems will be a highly desired skill (Thompson, 2016).

Collaborative problem solving is not only useful for completing job tasks; it is also applicable to maintaining a positive work environment. Managers may use collaborative problem solving to resolve issues among employees by engaging in collaborative discussions to reach a common understanding of the problem at hand and to negotiate a solution (Bernstein & Ablon, 2011).

Social and Emotional Intelligence
Emotional intelligence (EI), sometimes referred to as social and emotional intelligence, refers to an individual’s capacity to recognize one’s own and others’ emotions, use this knowledge to inform thinking and behavior, and adapt to meet goals. The concept has had a controversial history since the mid-1990s. Multiple definitions of EI—and various measures associated with each definition—exist today. Some studies have found positive correlations between EI scores and job performance and leadership skills; other studies find no unique contribution of EI to these outcomes beyond correlations accounted for by general intelligence and measures of generally accepted personality traits. We do not delve into the history and nuances of EI here, but instead summarize literature regarding EI’s perceived place among the skills needed in the workplace of the future.

Frey and Osborne (2013) conferred with experts in machine learning to determine the binary likelihood (i.e., yes/no) of automating 70 occupations based on their O*NET characteristics. These occupations were selected from the full suite of 702 O*NET detailed occupations based on confidence in the automation rating. Authors then used statistical modeling approaches to estimate the probability of automating the remainder of the occupations. After extensive analysis, the authors conclude “…as technology races ahead, low-skill workers will reallocate to tasks that are non-susceptible to computerization— i.e., tasks requiring creative and social intelligence. For workers to win the race, however, they will have to acquire creative and social skills” (p. 45).

Using a very different approach, PricewaterhouseCoopers (PwC; 2018) began a collaboration with the Said Business School in Oxford in 2007 to map influential business factors. The study
authors postulated four “worlds of work” to emerge by 2030 in which potential workplace scenarios are described in four quadrants defined by two dimensions: fragmentation vs. integration and collectivism vs. individualism. The authors developed descriptions of each scenario, including a timeline of milestones between 2020 and 2030, major characteristics of the quadrant, implications for workers, what the workforce will look like, and organizational challenges. Following a discussion of all four worlds, they predict the following about jobs: “Automation will not only alter the types of jobs available but their number and perceived value. By replacing workers doing routine, methodical tasks, machines can amplify the comparative advantage of those workers with problem-solving, leadership, EQ (Emotional Intelligence), empathy and creativity skills” (p. 30). PwC commissioned a survey of 10,000 individuals in China, India, Germany, the U.K., and the U.S. and found that 76 percent of respondents agreed or strongly agreed that they had emotional intelligence.

**Intrapersonal Skills**

**Time Management**
Time management skills encompass a variety of specific abilities: estimation of effort, scheduling, prioritizing, delegation, and monitoring a to-do list, among myriad others. An individual with strong time management skills can not only project the amount of time and effort a given task will require, but also inhabit the mindset to meet deadlines and, perhaps as importantly, recognize when a deadline cannot be met and adapt accordingly.

While employers have historically valued employees with solid time management skills, in the expanding gig economy the individual entrepreneur’s personal success depends upon it. When multiple gigs are underway, the ability to schedule and complete each gig as though it was the individual’s only job is critical. In a very dynamic environment where freelancers and hiring agencies are mixed-and-matched in various combinations for specific tasks, the entrepreneur cannot rely upon the understanding of a long-time employer who is familiar with the individual’s work and is sympathetic when projects fall behind. Task matching search engines such as Upwork (www.upwork.com), TaskRabbit (www.taskrabbit.com), or Gigwalk (www.gigwalk.com) collect customer satisfaction data and use this feedback to determine whether to match entrepreneurs to future tasks. Poor ratings due to a lack of effective time management could prevent further assignments.

**Efficiency**
Similar to time management skills, traditionally employers have valued efficient employees. For the 21st annual survey of CEOs worldwide, PwC interviewed 1,293 CEOs in 85 countries, including 104 from the United States, in October and November of 2017 (Ryan, Sapin, Rao, & Ampil, 2018), U.S. CEOs are hiring for broadly relevant digital skills and collaborative, creative, and efficient work styles.

In a gig economy, however, efficiency is particularly important. The individual entrepreneur may face challenges of scalability. At the extreme, the artisan who produces hand-made items can only produce so much; the individual service provider can only manage a limited number of clients or tasks. In order to scale up—which may be necessary in order to obtain a livable wage—the entrepreneur must be efficient.

43 A gig economy refers to a labor market characterized by the prevalence of short-term contracts or freelance work as opposed to permanent jobs.
Efficiency can be instantiated in a variety of ways in a gig environment. In some job markets, the individual entrepreneur may add staff and delegate work in order to increase production; however, in the examples just cited—creator of artisanal handmade items or tasker such as a personal shopper or errand runner—adding staff may not be feasible. Alternatively, the worker may use technology to offload mundane or repetitive tasks and free up time for more creative or complex work, requiring human skills. For example, subscribing to a task-matching search engine is an efficient way to seek work, relative to searching for opportunities and applying individually for each. Thirdly, a worker may leverage innovation or creativity to complete tasks more efficiently.

**Adaptability**

One of the common concerns about the future workplace is that automation will obviate the need for humans to perform large categories of jobs. Certainly, computerization and robots have been demonstrated to be effective replacements for humans in predictable, repetitive environments such as assembly line work. Further, AI has been successfully deployed to rapidly and accurately process large amounts of data to detect patterns and make complex, data-informed decisions. More recently, AI learning systems have been trained to determine optimal ways to conduct certain processes, and monitor their own ongoing effectiveness for further improvement. In this way, automated systems have the capacity for selected adaptability.

Many jobs, however, are less well-suited to automation. Autor and Dorn (2013) note that as computerization has become increasingly affordable, low-skill workers have shifted from routine tasks to the service industry. They contend that these service occupations are somewhat immunized against automation due to their reliance upon a combination of factors including direct physical proximity and flexible interpersonal communication.

In addition to adaptability being key to specific careers, adaptability will also be integral to the projected evolving career path an individual will undertake over the course of a working lifetime. As companies demand upskilling or reskilling, the adaptable employee will be at an advantage. McKinsey (2017) conducted interviews with experts from industry and academia for the April 2017 Digital Future of Work Summit in New York. Experts included professors and executives from NYU, C3 IoT, New America, WorkMarket, LinkedIn, Arena, and McKinsey Global Institute (MGI). MGI partner Susan Lund opined “For young people today, what’s clear is that they’re going to need to continue to learn throughout their lifetime. The idea that you get an education when you’re young and then you stop and you go and work for 40 or 50 years with that educational training and that’s it—that’s over. All of us are going to have to continue to adapt, get new skills, and possibly go back for different types of training and credentials. What’s very clear is that what our kids need to do is learn how to learn and become very flexible and adaptable” (p. 2).

Finally, in a gig economy the successful entrepreneur must be prepared to provide services to multiple employers in a variety of environments. Being able to adapt to technical and administrative requirements will serve the independent contractor well.

**Blended Skill Sets**

**Learning Agility**

Learning agility is a skill necessary for the development of skills of the future. It involves many of the skills required for the work of the future, including adaptability, tolerance for ambiguity, communication and listening skills (Eichinger & Lombardo, 2004), but applies them to the development of other skills. DeMeuse, Dai, and Hallenbeck (2010) indicate that learning agility
is related to past experience, self-awareness, and the ability to handle complexity. Mueller-Hanson, White, Dorsey, and Pulakos (2005) relate learning agility to adaptability.

**New Media**
New media refers to the emerging means of communication with large groups of people, and includes the internet, as well as more recent interactive, digital platforms through which more and more people access and consume information (Wynne, 2017). With increasing globalization and dispersion of the workforce, and with a growing number of employers creating and maintaining a new media presence, the ability to effectively navigate in this environment will become an increasingly valued skill (Gribben, Becker, & Dickinson, 2018).

Several core competencies have been identified as essential for participating in new media (Jenkins, Purushatma, Weigel, Clinton, & Robison, 2009). These include:

1. **Play** - the capacity to experiment with one’s surroundings as a means of problem-solving
2. **Performance** - the ability to adopt alternative identities to improvise or discover new things
3. **Simulation** - the ability to interpret and construct dynamic models of real world processes
4. **Appropriation** - the ability to sample and repurpose media content in meaningful ways
5. **Multitasking** - the ability to scan one’s environment and focus in on the key details
6. **Distributed cognition** - the ability to have meaningful interactions with tools that expand mental capacities (e.g., calculator, Wikipedia)
7. **Collective intelligence** - the ability to pool and compare knowledge with others to achieve a common goal
8. **Judgment** - the ability to assess how reliable and credible various information sources are
9. **Transmedia navigation** - the ability to follow stories and information across multiple media
10. **Network** - the ability to search for, synthesize, and distribute information
11. **Negotiation** - the ability to travel across diverse communities, recognizing and respecting differing perspectives, and understanding and following alternative norms
12. **Visualization** - the ability to translate information into visual models and to understand the information that other visual models are conveying

**Projections of Shifts in Future Skills**

**Skills Expected to Increase in Demand**

Literature on trends in workforce skills typically does not include information to determine the rate of the trend. Thus, it is difficult to identify which skills are expected to increase in demand. In the next sections, we present skills that have been deemed critical for the future workforce and show evidence of expected increase in demand. Based on projections of work in the future, there is evidence of expected increase in jobs demanding the following subset of future postsecondary skills.

**Complex Problem Solving**

Futurist Cynthia Wagner (2011) forecasts a limitless supply of future problems to solve. Educators consider critical thinking and complex problem solving a core skill for students (Lara, 2018; P21, 2016) in part because employers increasingly seek workers with complex problem-solving abilities.
solving skills. Employers have consistently rated complex problem solving as the most essential competency for career preparedness (National Association of Colleges and Employers, 2017). The ability to address complex problems is particularly important in work environments characterized by rapid change (Middleton, 2002).

**Computational Thinking**

Opportunities to learn how to code, especially for girls (see girlswhocode.com) amidst efforts to achieve gender parity in the computer science industry, have exploded in popularity (Bourque, 2016). Games that teach programming logic are being marketed for children as young as four-years-old (see Kodable.com). Hackathons offer monetary awards and bragging rights, spurring innovative design and problem solving. However, Grover (2018) argues that not everyone will become a computer programmer. Computational thinking – the “ability to translate vast amounts of data into abstract concepts and to understand data-based reasoning” (Davies, Fidler, & Gorbis, 2011, p. 10) – should be taught to every student (Grover, 2018).

**Communication**

Using government job-growth projections, the Pew Research Center identified the fastest growing occupations and skills and preparation requirements for working in those fields (DeSilver, 2016). Nine of the top 10 fastest growing occupations, with projected growth of 5.2% to 13.1% in the ten-year period 2014 to 2024, require job preparation (i.e., formal education, on-the-job training, and prior related experience), interpersonal skills (e.g., communication), or both.

**Cultural Sensitivity and Communication**

According to Davies, Fidler, and Gorbis (2011), ease and sensitivity in working with culturally diverse colleagues will become an important skill for all workers, not just those who work in global corporations. Employers are increasingly recognizing the value of cultural competence and communication skills among new hires (Vozza, 2016), especially when those skills are needed to perform future jobs that involve interaction on a global scale.

**Collaborative Problem Solving**

The demand for collaborative problem-solving skills is anticipated to experience high levels of growth in the future (Thompson, 2016). In recent years, employers have consistently rated collaboration among the top competencies needed for career preparedness, and recognize it as the top attribute for setting apart a potential employee’s resume (National Association of Colleges and Employers, 2017). Industry strategists expect collaboration to be a top business objective in the workplace of the future (Bowles, 2018; P21, 2016). Virtual collaboration (Davies, Fidler, & Gorbis, 2011) will be more in demand as multifaceted problems are too complex to be solved within one discipline or organization (e.g., climate change) and necessitate working on virtual teams. Working on collaborative teams allows for skills gaps to be bridged and increases efficiency (Boyer, 2017).

**Social and Emotional Intelligence**

Davies, Fidler, and Gorbis (2011) declared social and emotional intelligence as one of ten skills critical for success in the future workforce. The Partnership for 21st Century Learning (2016) includes interpersonal skills as a key skill for students to learn before graduation from high school and embarking on postsecondary pathways.

**Adaptability**

Business trends indicate a need for more employees who are comfortable adapting to frequent changes to their work and the workplace (Davies, Fidler, & Gorbis, 2011; McKinsey & Company,
As organizations are required to respond quickly to changes in an increasingly globalized and technologically advanced world, employers will seek an agile workforce capable of responding to unanticipated change with speed (Breu, Hemingway, Strathern & Bridger, 2002). Workers of the future may be expected to rotate among a variety of roles and tasks (Wadors, 2017). With increasing contract and gig positions expected in the future (Yaraghi & Ravi, 2016), individuals will rely on their adaptability skills as they move from one job to another. Additionally, adaptability skills will facilitate keeping up with changes in technology. Adaptive thinking is a key to innovation and creative problem solving (Davies, Fidler, & Gorbis, 2011).

**New Skills Expected to Emerge**

While information about specific future jobs was scarce, prediction of new skills is non-existent. Using projections of new jobs requiring creativity and social and emotional intelligence, we can extrapolate the need for skills drawing on creativity and social and emotional skills. Increases in artificial intelligence (AI) applications may lead to a need for new skills combining knowledge of AI and other skills to further the research, development, and integration of AI into the workplace.

When futurists discuss workplace skills they tend to focus on combinations of current skills, such as skills in multiple subject-area domains (see transdisciplinarity in Davies, Fidler, & Gorbis, 2011) or skill in cognitive and non-cognitive areas (e.g., analytical and interpersonal skills). By 2030, some new jobs will likely demand completely new skills that we cannot yet imagine and therefore currently cannot describe or label.

**Summary of Themes of Skills for the Future**

When it comes to skills for the future, the types of jobs offered and changes to workplace processes will be the major drivers of skills needed to enter the workforce. By extension, skills for jobs needing postsecondary education or training require skills for success in school as well as job-specific and cross-career skills. Although change is expected in the world of work, the amount and direction of change is unknown. However, this review of relevant literature points to some overarching themes that provide a solid base for making predictions about career skills that students who entered kindergarten in 2017 will need in 2030 when they graduate from high school. Careers will likely look very different from those their parents were prepared for, particularly in terms of the number of jobs and variety of skills needed across and within those jobs over their lifetime.

Jobs of the future like those of today will require a mix of cognitive, interpersonal, and intrapersonal skills. The specific set of skills and which ones are most important will vary depending on the pathway a student follows. Increasingly, blended skill sets (e.g., learning agility and new media) will be needed. Fields that had previously been quite separate may be blended in new ways, requiring combinations of skills not seen before. Existing jobs may be blended with new technologies to create positions we have never seen (think: space junk recyclers) and requiring new skills or blended skill sets.

Industry believes “data is the new oil.” To meet the expected rise in demand for employees to work with data, employers will seek individuals with skills related to manipulating, analyzing, interpreting, and illustrating data.

High school graduates of 2030 will set out on career pathways characterized by change. Whether they work independently through the gig economy, or move among multiple employers or across multiple departments or projects, workers of the future will likely need to be adaptable and agile as they will be asked to respond quickly to unanticipated changes. They will need to
draw upon cross-cultural and virtual collaboration skills as they find themselves part of an increasingly diverse and dispersed workforce. Workers will become lifelong learners as jobs continually evolve to meet changing demands and to incorporate the latest innovations.

With some sense of what the future holds, a key next step to ensuring that students graduate high school in 2030 prepared for success in their postsecondary pathways is to measure the skills needed to perform the jobs of the future. Assessing future postsecondary skills provides a metric for understanding and monitoring how prepared the next generation is for life after high school.
References


Land, M. H. (2013). Full STEAM ahead: The benefits of integrating the arts into STEM. Procedia Computer Science, 20, 547-552


Thompson, C. (2016, January 21). The top 10 skills that will be in demand by all employers by 2020. Business Insider.


Appendix J. Literature Review: Skills of the Future


A Review of State Indicators of College and Career Preparedness

Introduction

The purpose of this memorandum is to review states’ practices for measuring college and career readiness. The memorandum begins by exploring how states have defined the constructs of “college and career readiness.” Next, common and well-established measures such as college entrance exams and statewide assessments in English Language Arts (ELA) and mathematics are described. Then, some less-established academic indicators are discussed. Finally, the memorandum discusses other career or work readiness indicators, including non-academic and other innovative indicators, that states may be including in their suite of exams to tap into career or workplace skills.

The information included in this memorandum was obtained from (a) states’ federal peer review submissions, (b) publicly available information on websites (e.g., state websites, research organization websites, test vendor websites, governmental organization websites), and (c) a review of current articles and scientific literature (e.g., EdWeek articles). Each listed indicator/assessment includes its own research base and documentation; it is beyond the scope of this memorandum to include all of that information. However, enough detail is provided to allow interested parties to discover more information about assessments of interest.

As this memorandum will illustrate, assessments specifically designed to address “career readiness,” including non-cognitive, social-emotional learning, and career exploration, are not widely used by states. Some of the indicators described in herein may be used by a single state or may not be administered to all students or even to all schools/districts within a state. However, many states are in the midst of procuring new statewide assessment contracts and the landscape of assessment is ever-changing. Thus, there are likely additional indicators of preparedness for college and career that are not included in this memorandum. As such, this memorandum represents a preliminary effort to catalog some of the innovative ways that states are beginning to approach college and career preparedness.

Framing the Need to Address Readiness for College and Career

Preparing students for post-graduation opportunities has long been a priority for states, districts, schools, teachers, and parents. The current expansion of the global economy and ongoing labor market shifts has refocused attention on the readiness requirements for student success in an ever-changing post-secondary landscape. The current workforce demands employees who can fill jobs that require a combination of technical and academic skills (Carnevale, Smith, & Strohl, 2010). However, employers struggle to find such qualified workers (Organization for Economic Co-operation and Development, 2013). Educational institutions struggle with defining and measuring the knowledge, skills, abilities, and dispositions necessary for success in the current post-secondary world. These efforts are key to guiding educators’ efforts to identify which students are on track to succeed as they graduate from high school.

Porter (2018) provided an overview of the current landscape of occupations in the United States and their requirements during a talk at the National Conference on Student Assessment (NCSA). He described ways that states have attempted to upgrade their tests to prioritize the most relevant knowledge and skills for students. The summary figure is reproduced below.
According to Mishkind (2014), 37 states defined college and career readiness as of 2014. For 33 of those, a single definition was used to describe both college and career readiness. The rationales for using a single definition tend to include views on the interconnectedness and similarities between visions of preparedness for college and preparedness for career. For example, the Oregon Investment Education Board\(^44\) indicates that, “the overarching skills and strategies required for students of all ages entering colleges and careers are consistent.” New Hampshire, in their ESEA flexibility request\(^45\) stated, “Evidence and experience indicate the knowledge and skills needed to succeed in college and career are greatly similar, and that all graduates will need some form of postsecondary education or training to succeed during their careers.”

The U.S. Department of Education (ED) currently indicates that “Nearly every state now has adopted college- and career-ready standards.” This quote refers to college and career readiness specifically in the subjects of mathematics and ELA. ED also describes the requirements for approval of state assessment and accountability systems under ESSA related to standards, indicating that “All states approved for ESSA flexibility have engaged in one of the following endeavors to raise expectations for students’ academic performance:

- Upgraded their existing standards to make them more rigorous by working with their four-year public universities to certify that mastery of standards ensures that students

---


will not need to take remedial coursework upon admission to a postsecondary institution in the system; or

- Adopted and implemented common standards developed by a consortium of states that build toward college and career readiness.\(^{46}\)

According to ED, as of May 31, 2018, 33 states (plus the District of Columbia) included a measure of college and/or career readiness in their ESSA plans.\(^{47}\)

In states where there are separate definitions, there is a tendency for considerable overlap in the definitions. For example, the Nebraska Department of Education defines a career ready person as follows, "A career ready person capitalizes on personal strengths, talents, education and experiences to bring value to the workplace and the community through his/her performance, skill, diligence, ethics, and responsible behavior.\(^{48}\) The Nebraska Career Readiness Standards indicate that when students are career ready, they are prepared for the next steps in their lives—whether that means getting their first job or beginning their college ‘career’ (which eventually leads to the workplace as well). Being career ready also means being ready for life." Nebraska is one of an increasing number of states to also include career and technical education as part of their statewide accountability system. They provide programs outside of the regular school day for students to connect with resources to develop interests and skills for future success.

Mishkind (2014) identifies specific components of the states’ definitions of college and career ready and places them into the following six categories:

- Academic knowledge (e.g., English, mathematics, core subjects, preparation to take credit-bearing courses in related subjects)
- Critical thinking and/or problem solving (e.g., reasoning, problem solving, analysis, evaluation, presentation of ideas and information)
- Social and emotional learning, collaboration, and/or communication
- Grit/resilience/perseverance
- Citizenship and/or community involvement
- Other additional activities

Distribution of these components among the 37 states identified by Mishkind (2014) is presented in Figure 2. Similar reports reflecting the state definitions under ESSA have not yet been published.

This memorandum treats “readiness” and “preparedness” synonymously, although states may make distinctions between them. Treatment of the terms is not consistent from state to state, with some referencing “readiness” only as a link to college placement tests’ established readiness benchmarks, while others treat readiness more generally. This memorandum is concerned with documenting the ways that states address readiness, so both terms were used in its preparation and no distinctions were made between them.

---

\(^{46}\) See [https://www.ed.gov/k-12reforms/standards.](https://www.ed.gov/k-12reforms/standards)

\(^{47}\) See [https://www.ecs.org/50-state-comparison-states-school-accountability-systems/](https://www.ecs.org/50-state-comparison-states-school-accountability-systems/)

\(^{48}\) Definition adopted in 2010, see [https://www.education.ne.gov/nce/careerreadinessstandards/](https://www.education.ne.gov/nce/careerreadinessstandards/)
Figure 2. Components of state definitions of “College and Career Ready”

Typical Measures of Readiness

While states are exploring expanded definitions of readiness, their federal peer review submissions indicate that, for the most part, academic or college readiness is the aspect that is measured. Academic readiness is typically measured by traditional assessments, and work or career readiness is considered either synonymous or is described as a set of traits or characteristics that extends beyond academic readiness. States assume academic readiness is a necessary component of preparation for both college and career. Most states do not include measures of non-academic readiness in their accountability systems.\(^49\)

1. State Accountability Tests

State are required under ESSA to test students in grades 3-8 and high school in ELA and mathematics. Some statewide assessments are administered on computers while others are administered via paper and pencil. These assessments are typically a mix of multiple-choice and short constructed response (students provide a brief written response) items. Some state tests also include multiple-select items (which are selected response items with potentially more than one correct answer), essay or extended constructed response items (longer writing pieces), or technology enhanced items (TEIs). TEIs must be computer administered and may include matching, drag-and-drop, ordering, or other tasks that are more readily performed in the computer administered environment.

\(^{49}\) See https://www2.ed.gov/admins/lead/account/stateplan17/statesubmission.html, for a full description of states’ peer review submissions.
Most statewide summative assessments are based on the Common Core State Standards for (CCSS) ELA and mathematics. The figure below, from the Association for Supervision and Curriculum Development (ASCD),\(^{50}\) shows the state level adoption of the CCSS as of 2018. States highlighted in green are adopters, Minnesota, (highlighted in blue) adopted the ELA but not mathematics standards. States highlighted in gray did not adopt the CCSS. This graphic shows there is more agreement among states than not regarding what academic content students should learn in school, at least for ELA and mathematics.

![Figure 3. CCSS adoption, by state.](http://www.ascd.org/common-core-state-standards/common-core-state-standards-adoption-map.aspx)

In addition to largely agreeing on tested content, states have been steadily moving toward asking students to perform more complex and challenging tasks on summative tests. The testing consortia, Smarter Balanced Assessment Consortium and the Partnership for Assessment of Readiness for College and Career (PARCC), require a great deal of strategic thinking (critical thinking, reasoning, and developing a plan) and even extended thinking (thinking to investigate a problem and synthesize information) in a substantial portion of their assessment items. Figure 4 compiles information on the level of complexity among several summative state tests (Porter, 2018). This research was compiled from studies conducted by HumRRO (Schultz, Wiley, Michaels, & Dvorak, 2016), the Thomas B. Fordham Institute (Doorey & Polikoff, 2016) and Rand (Faxon-Mills, Hamilton, Rudnick, & Stecher, 2014).

---

States have adopted challenging content standards and state assessments have become increasingly more complex. However, comparability of students’ scores across states remains challenging. Unless states use a common assessment with common proficiency definitions (such as those states in consortia), students’ proficiency may depend on the state where they were tested. To address comparability, states participate in reading and mathematics testing on the National Assessment of Educational progress (NAEP). NAEP is a complex suite of assessments based on rigorous academic frameworks documents. NAEP tests do not report individual student scores, but report at the state level based on samples of students. The National Center for Education Statistics (NCES) produces maps between state and NAEP proficiency standards, allowing states to compare the rigor of their standards using a common assessment. States often reference the NAEP Mathematics and Reading Frameworks and the NAEP proficiency standards when they design their own assessments.

2. High School GPA

Student academic performance is also tracked by grade point average (GPA), a simple transformation of the letter grades they receive for their courses. GPA is not typically tracked at the state level as part of accountability, but grades are commonly used as an indicator of college readiness and GPA is a required component on most college applications. Some states, notably Kentucky, use grades to determine supplemental funding provided to help students pay for college. The Kentucky Educational Excellence Scholarship (KEES) program rewards students who graduate with at least a 2.5 GPA, and the amount of money provided to students goes up if their GPA is higher.

The measure of GPA includes more than academic knowledge (Farrington, Roderick, Allensworth, Nagaoka, Keyes, Johnson, & Beechum, 2012). The requirements for achieving strong grades in classes also incorporates perseverance, compliance, and time management. Depending on the nature of their classes, students may also need to demonstrate their ability to

---

51 See [https://nces.ed.gov/nationsreportcard/](https://nces.ed.gov/nationsreportcard/).
52 See [https://nces.ed.gov/nationsreportcard/studies/statemapping/#](https://nces.ed.gov/nationsreportcard/studies/statemapping/#).
53 See [https://www.kheaa.com/website/kheaa/kees?main=1](https://www.kheaa.com/website/kheaa/kees?main=1). Students can increase their KEES scholarships with high test scores as well (e.g. ACT, Advanced Placement, International Baccalaureate, Cambridge Advanced International).
collaborate with others, communicate their ideas, and even to construct products. Grades are more far-reaching than summative tests. However, grades are not standardized from teacher to teacher, school to school, or district to district, and they are not typically used by states in their accountability computations.

3. Course Requirements/Completion

States require that all students complete a certain number of courses in specific subjects prior to graduation. These required courses include ELA, mathematics, science, and social studies; however, they may also include arts, health, physical education, foreign language, or others. The Education Commission of the States (ECS) maintains a website that details all states course requirements for graduation. In addition to state requirements, districts may add their own requirements for earning a high school diploma.

4. End-of-Course Assessments

States may use either summative tests for high school students administered in particular grades, or they may tie testing to the completion of courses, or both. When states choose to use end-of-course (EOC) exams, they are most often tied to English 10 or 11, and to either Algebra 1 or Algebra 2 courses. For example, PARCC created assessments for multiple English and mathematics courses for states, including subject-specific mathematics courses like Algebra and Geometry, as well as Integrated Mathematics. These types of tests limit the amount of content to be assessed to that addressed by a single course (as opposed to the full range of high school standards), but they assume all courses with a common name cover essentially the same content. States may require or encourage schools to use the assessment scores as part of student’s grades, or as a course completion requirement. EOC assessments may be used for state accountability computations.

Even states who use EOC exams vary greatly on the number of EOC exams offered, the courses for which they are offered, and whether EOC exams are used for accountability. For example, a state may use an Algebra 2 EOC exam for high school accountability, but offer an Algebra 1 EOC exam that is not used. For states with science accountability in place for high school, this is further complicated by course taking patterns (science courses do not follow a particular sequence and students may not take certain science courses). Some states provide EOC exams for Biology for accountability, and require that all students take a biology course, while others require a more general science exam administered at a particular grade level. Others require the Biology EOC exam, but supplement it with other science content to address physical and earth/space sciences.

5. Graduation Exit Exams

The following thirteen states used high school graduation (or exit) exams as of 2017; Florida, Indiana, Louisiana, Maryland, Massachusetts, Mississippi, New Jersey, New Mexico, New York, Ohio, Texas, Virginia, and Washington. Exit exams require that students demonstrate some minimum level of academic competency, typically in ELA and mathematics, as a condition of graduating. Students who do not pass the exams do not receive a high school diploma. While not specifically designed to indicate college or career readiness, high school exit exams help certify to employers that graduates can be expected to have a minimum level of competency in

---

54 See www.ecs.org
55 See https://www.fairtest.org/graduation-test-update-states-recently-eliminated
the tested subjects. These types of exams are controversial for several reasons, including the argument that they may reduce graduation rates and increase dropout rates, and they may test content that is commonly found in middle school standards rather than in high school. Exit exams tend to focus on literacy and numeracy, rather than the more complex high school standards. The literature is mixed on these topics and too extensive to adequately review here. Arkansas, Arizona, Georgia, Idaho, Minnesota, Oklahoma, Rhode Island, and South Carolina had graduation exams, but they recently terminated the requirement for high school graduation. California, Nevada, and Pennsylvania currently have a moratorium on graduation tests.

6. College Entrance Exams

College entrance exams are another commonly used type of assessment of readiness for specific courses in colleges or universities. Common examples of entrance exams include College Board’s Accuplacer® assessments, ACT’s Engage® assessments (ACT previously promoted a placement test called Compass), and a multitude of college- and university-created placement tests. While not typically used statewide, these assessments provide indications for higher education of the likelihood that students will successfully complete specific courses. Results from these assessments may be used to place students into non-credit remedial courses at a college or university.

In the early years of its reporting as part of state accountability systems, college entrance exams were frequently used as an indicator of college and career readiness. For example, the Southern Regional Education Board (SREB) presented changes in average SAT composite scores as an indication of trends in college and career readiness (SREB, 2012). As the concept of readiness for college and career has become better articulated, it is clear college entrance exams may not be the most appropriate tool for measuring career readiness.

The ACT and SAT programs are very clear about their intended purposes. The ACT is designed to measure “academic readiness for college.” Similarly, the SAT measures what students “need to succeed in college.”

Approximately half of states’ accountability systems incorporate college entrance exams as an indicator of college and career readiness. Typically, student performance (e.g., meeting established benchmarks) is measured, but in a small number of states only rates of participation in these exams is reported. In no states are college entrance exams the sole indicator of college and career preparedness. Typically, these scores are reported along with several other indicators, such as performance in advanced coursework (e.g., AP or IB classes), as well as performance on WorkKeys®, earning credentials or certifications, and/or performance in career and technical education classes (American Institutes for Research, 2018).

College entrance exams have the benefit of large, well-established bodies of validity evidence, and as such may be viewed as a cost-effective resource for measuring postsecondary preparedness. However, one potential concern is the extent to which the content knowledge and skills measured by these exams aligns with the content standards adopted in accordance with state law. Most states have adopted standards that are closely related to the CCSS. ACT

---

56 See https://accuplacer.collegeboard.org/
(2010) conducted an evaluation of the alignment between its College Readiness Standards and the CCSS and found substantial levels of alignment in terms of reading and mathematics, with weaker alignment in writing. Similarly, the College Board found strong alignment between the SAT and common core anchor standards in reading, writing, language, and mathematics (Vasavada, Carman, Hart, & Luisier, 2010). However, a recent study by Assessment Solutions Group found it would not be appropriate to allow school districts in Florida to opt to use the ACT or SAT in lieu of its state assessment (Roeber, Olson, & Topol, 2018). Currently, 25 states require all high school students to take the ACT or SAT, and 12 states use the ACT or SAT for accountability.61

College entrance exams such as ACT and SAT identify cut scores, or benchmarks, to assist in defining and evaluating student progress toward, and achievement of, adequate levels of college preparedness. States may be required by law to identify benchmarks specific to their own state. For example, the Wisconsin Department of Public Instruction convened panels of educators and other stakeholders to evaluate ACT benchmarks and determine the appropriateness of identifying state-specific benchmarks (Wisconsin Department of Public Instruction, 2015).

**Academic Readiness Measures Beyond ELA and Mathematics Achievement**

Academic preparation for college and career involves more than preparation in ELA and mathematics. High school students take at least 20 courses in grades 9–12, and only eight of those courses are typically required to be English or mathematics courses.62 Students also take courses in science and social studies. They may take classes in the arts, technical courses, foreign language courses, and others. The NCLB and ESSA helped create a focus on English and mathematics, but now states have largely adopted the Next Generation Science Standards (or similar state-specific, three-dimensional science standards) and many have also adopted social studies standards. Many states have begun to test these subjects as part of their accountability systems.

1. **Science Assessments**

Nearly two-thirds of U.S. students live in states that have education standards influenced by the Framework for K–12 Science Education63 and/or the Next Generation Science Standards (NGSS)64. Figure 5 shows these states in blue and yellow (blue states have adopted NGSS, yellow adopted their own standards based on National Research Council [NRC] frameworks). Both documents promote a more complex phenomenon-based view of science and stress students’ capacity to use their science knowledge in unfamiliar contexts. They address science and engineering principles, cross-cutting concepts, and content knowledge and expect those ideas to intersect for students through rich science instruction.

61 For a complete list of states requiring ACT or SAT see https://www.edweek.org/ew/section/multimedia/what-tests-does-each-state-require.html.
62 See www.ecs.org
63 See https://www.nap.edu/read/13165/chapter/1.
64 See https://www.nextgenscience.org/.
Nineteen states and the District of Columbia (representing over 36% of U.S. students) have adopted the Next Generation Science Standards (NGSS). The 19 states are Arkansas, California, Connecticut, Delaware, Hawaii, Illinois, Iowa, Kansas, Kentucky, Maryland, Michigan, Nevada, New Hampshire, New Jersey, New Mexico, Oregon, Rhode Island, Vermont and Washington.

Twenty states (representing 29% of U.S. students) have developed their own standards based on recommendations in the NRC Framework for K-12 Science Education. The 20 states are Alabama, Colorado, Georgia, Idaho, Indiana, Louisiana, Massachusetts, Mississippi, Missouri, Montana, Nebraska, New York, Oklahoma, South Carolina, South Dakota, Tennessee, Utah, West Virginia, Wisconsin, and Wyoming.

Figure 5. State science standards summary.

Most states test science once in elementary school (usually at grade 4 or 5), once in middle school (usually in grade 7 or 8), and once in high school (usually in grade 11 or as an end-of-course biology assessment). Science assessments are only now catching up to the adopted science standards. Some states (e.g., Kentucky) have implemented highly complex phenomenon-based assessments for science (similar to scenario-based tasks used in NAEP). These assessments are similar to passage-based ELA assessments, where students are given a passage to read and then answer questions requiring them to interpret or interact with the passage. In science, students would be given a thorough description of some scientific phenomenon, which may include experimental results, data tables, graphs or charts, or other information referencing the problem the students are expected to help solve. Students demonstrate their understanding of science content by interacting with the phenomenon. Phenomenon-based assessments can take longer to develop, may be more expensive than more traditional tests, and may take longer for students to complete. States are struggling with
the structure of science assessments, how to ensure alignment to complex three-dimensional standards, and how best to report science scores for students, as well as for schools.

2. Social Studies Assessments

Historically, states have not adopted common social studies standards or tests of social studies. While there is a growing trend for states to adopt social studies standards (see above), assessments of social studies remain less common than ELA, mathematics, or science. Social studies assessments, when administered, are typically given at the same grade levels as science. The content of state-wide social studies tests is based on standards that vary considerably from state to state. Some states use EOC assessments of US History in high school, while others test social studies more broadly at specific grade levels.

The National Council for Social Studies (NCSS) has published social studies standards, but they have not been widely adopted by states or included as the basis for assessments in state accountability systems. The NCSS has identified five disciplines of social studies: history, geography, civics and government, economics, and psychology. Those disciplines address ten themes of social studies, including:

- Culture
- Time, continuity, and change
- People, places, and environments
- Individual development and identity
- Individuals, groups, and institutions
- Power, authority, and governance
- Production, distribution, and consumption
- Science, technology, and society
- Global connections
- Civic ideals and practices

Career or Work Readiness

There is growing acceptance of the notion that readiness for college and readiness for career are distinct but overlapping constructs. Patelis (2018) notes there are sets of skills at the intersection of career and college preparedness. Similarly, ACT has developed an integrated framework that defines the core academic skills for readiness from kindergarten through one’s career. The common thread here is that there are key academic components that are common to postsecondary readiness in general, whether a student opts to continue learning in a postsecondary educational setting or in a more applied, workplace-based setting.

---

Career preparedness includes three major skills needed to perform work: (a) core academic skills; (b) employability skills; and (c) technical, job-specific skills (ACTE, 2010). Academic skills include foundational academic knowledge in mathematics, ELA, and science or technology, including the application of these skills, which were described in the prior section of this memo. Employability skills are skills that are critical to workplace success and include critical thinking, adaptability, problem solving, oral and written communication, collaboration and teamwork, creativity, responsibility, professionalism, ethics, and technology use. Technical skills represent job-specific knowledge and skills needed to enter a career.

Next, some indicators of career readiness that are being used or being considered by states and districts are presented. Because many states are currently exploring career readiness indicators, due at least in part to expanded flexibility in ESSA for using such measures, there is not clear documentation currently available on which districts within which states are using these indicators.

1. Assessments of Career or Work Readiness

The Center on Education Policy (2013) found via survey that 36 of 46 participating states assessed students on academic skills related to career readiness. Smaller numbers reported assessing students on technical skills (33 states) and employability skills (28 states). They found the tests most commonly used to assess employability skills and/or applied academic skills were the Armed Services Vocational Aptitude Battery (ASVAB) and WorkKeys® assessments. The ASVAB tests are designed to measure verbal, mathematics, science and technical, and spatial domains. WorkKeys® “measure foundational skills required for success in the workplace” via several assessments including applied mathematics, graphic literacy, workplace documents, applied technology, and business writing, among others.66

A smaller number of states reported administering the National Occupational Competency Testing Institute (NOCTI) assessments. NOCTI offers a Workplace Readiness assessment to assess employability skills at the high school level, as well as tests of employability skills that are targeted toward middle school students.67 Currently, 11 states use NOCTI in some capacity, but their use of the assessments is varied. Some require full census testing in high school, while others make the NOCTI assessments available for school or district use.

In fall 2017, the Center for Educational Testing and Evaluation (CETE) conducted its first operational administration of the Career Pathways Assessment System (cPass® assessments), which are designed to measure “both the knowledge and skills needed for specific career pathways.”68 In addition to a general Career and Technical Education (CTE) assessment, cPass® offers assessments of Comprehensive Agriculture; Power, Structural, and Technical Systems; Animal Systems; Plant Systems; Horticulture; Manufacturing Production; Comprehensive Business; Finance; Accounting; and Drafting. Currently, Kansas and Colorado are members of the collaborative that oversees administration of the cPass®.

Still other states have developed their own career readiness assessments. For example, Kentucky developed the Kentucky Occupational Skills Standards Assessment (KOSSA), which

---

68 See https://careerpathways.us/.
assesses a combination of academic, employability, and occupational skills specific to each of more than 30 career pathways.  

2. **Industry Certifications**

Achieve, Inc. found substantial variability of states’ approaches to incorporating industry certification into educational policy and processes (Muller & Beatty, 2009). Some states guide students along a CTE pathway, ensuring they complete the course requirements that will best prepare them for achieving certification within their preferred industry. Other states allow districts to develop approaches that meet their specific needs. Among the more innovative programs are the tiered diploma system developed in Florida that offers a credential-based graduation option and the “Core 40 with technical honors” diploma offered in Indiana.

3. **Descriptors of Traits or Characteristics**

While not written into statute or as part of the accountability computations, some states have published traits or characteristics of college and career ready students that go beyond academic expectations. These traits or characteristics may include citizenship, financial responsibility, goal setting, problem solving, and reasoning, among others. For example, in 2013, Michigan produced a list of *Characteristics of Career and College Ready Students*. Michigan’s source documents include the Michigan Literary Standards, National Research Council (NRC) Science and Engineering Practices, Michigan Social Studies Grade Level Content Expectations (GLCE) and High School Content Expectations (HSCE), Michigan Mathematics Practices, Michigan Visual, Performing and Applied Arts (VAPAA) Guidelines, and Career and Technical Education (CTE) Career Ready Practices. Under each source document, Michigan then lists the specific characteristics gathered from each source, grouped into four main headings (a) Technology and Tools, (b) Argument and Reasoning, (c) Communication and Collaboration, and (d) Problem Solving. Other states have addressed attributes of readiness in similar ways (e.g., Kentucky, Nebraska). According to ED, six states plan to use an art access/participation, or well-rounded education measure, in their accountability systems. States’ plans tend to focus on the school level, measuring access or participation rates, rather than on individual student's preparedness.

4. **Social and Emotional Skills Assessment**

Social and emotional skills assessment is a relatively new field for estimating students’ academic preparedness. The ACT Tessera is focused on the holistic growth of the “whole child.” It measures students’ social and emotional learning skills and provides data to help schools support goal setting and promote continuous improvement. ACT Tessera includes indicators of grit, teamwork, resilience, curiosity, leadership, and school climate. It uses a combination of self-report and forced choice and situational judgment test items to ensure accurate reporting and minimize student “faking.” Reports are provided at the student and school levels, and both students and schools receive recommended resources and strategies to address areas requiring increased focus.

---

69 See [https://education.ky.gov/CTE/kossa/Pages/default.aspx](https://education.ky.gov/CTE/kossa/Pages/default.aspx)


71 See [https://www.ecs.org/50-state-comparison-states-school-accountability-systems/](https://www.ecs.org/50-state-comparison-states-school-accountability-systems/)

5. Service Learning/Applied Experience

Service learning, for which students engage in projects or activities to improve their communities or address social problems where they can apply their learning, is increasingly reported as a requirement for high school graduation. To date, only Maryland and Washington D.C. have statewide service learning requirements; however, 19 other states allow districts to require service learning as a graduation requirement. A full description of service learning requirements for graduation by state can be found at the Education Commission for the States website (http://ecs.force.com/mbdata/mbquest3RTE?Rep=SL1301).

6. Interest or Suitability Inventories (Career Exploration)

Although they do not require interest or suitability inventories for career exploration, many states indicate these tools are available and they may provide links to them for interested students. Their use may also be promoted at the district or school levels. Some of these tools provide information at the classroom or school level, but the information provided varies greatly from one tool to the next. In addition, many colleges and universities encourage students to complete interest inventories or surveys prior to applying for admission or choosing a major. They may promote this activity through interaction with high school guidance officers or others with access to high school students nearing graduation.

When available, survey results are used to help students explore the fit between their interests or traits and potential future jobs or college major fields of study. These types of surveys have been used for a long time and there are a wide variety currently available to students. Some of the oldest are based on personality traits, with the personality traits matched to specific careers or job categories. Examples include the Myers-Briggs Type Indicator (MBTI) (Myers & McCaulley, 1985), Keirsey Temperament Sorter (1998), and Strong Interest Inventory (Donnay, 1997). These indicators of traits and their relation to jobs have been used for more than 50 years, with the Strong Interest Inventory used since 1927. The Five Factor Model (FFM) of personality traits is more commonly used today in business and industry, and it has been linked to academic performance indicated by grade-point-average (GPA) in high school and college (Poropat, 2009). These types of inventories assume the traits or interests are not (or are very minimally) mutable. They focus on using information about the students to match them with careers or job families that are best suited to their specific characteristics.

Below are some specific interest inventories/platforms that are currently in use.

6a. ASVAB Career Exploration Program (CEP)

A comprehensive career planning program, the ASVAB Career Exploration Program (CEP) is the only career planning resource available to high schools nationwide that offers high school students the opportunity to explore all paths to career satisfaction—education, work, training, military, and/or certification—because the ASVAB CEP assesses academic ability and career interests. The ASVAB CEP first helps students identify skill strengths based on their ASVAB multiple-aptitude skills test results and then links those skill strengths with an interest inventory that highlights work-related interest areas where students are most likely to succeed. ASVAB’s Find Your Interests Inventory is based on Holland’s occupational codes: Realistic, Investigative, Artistic, Social, Enterprising, and Conventional (RIASEC). The Holland Codes refer to a theory of career choice based on personality type. Finally, students access a catalog of careers with

---

73 See https://www.asvabprogram.com/
74 See https://www.hollandcodes.com/
relevant occupational data and a full suite of future-oriented planning tools that help them develop an action plan to share with parents and educators.

Equipped with an understanding of their skills and interests, students can explore careers that coincide with their skill-interest profile and their aptitudes. Participation is voluntary. Students in grades 10–12 and college are encouraged to use the ASVAB CEP to identify their skill strengths and explore potential careers that require those skills.

Students gain access to career exploration tools after they receive their scores. Then, they can independently explore any careers that interest them, including those found in the military. The Department of Defense (DoD) sponsors ASVAB CEP with a two-part mission: to provide a career exploration service to U.S. youth and to provide qualified leads to military recruiters. Participants have no obligation to military service, but 11th–12th grade students who choose can use their scores to explore enlistment. ASVAB CEP is provided to students at no cost. Currently, 12 states use military testing or enlistment as an indicator for accountability in their ESSA plans. Kentucky, for example, provides students multiple options for demonstrating postsecondary readiness, including a military readiness route that includes ASVAB and military enlistment.

6b. Career Cruising
Career Cruising’s (https://public.careercruising.com/en/) online interface engages students in exploring their interests and aspirations through interactive games starting in K–2, followed by career activities and assessments in later grades. Career Cruising connects the real world to the classroom with current career and labor market information, salaries, and educational pathways. With this resource, students can explore skills like financial literacy and goal-setting through video interviews, role-playing activities, and more. Using Career Cruising for K–12, students discover how their strengths and interests align with careers that inspire them, and build the academic plan to get them there. Career Cruising helps students understand how the schoolwork they do today impacts the life they live beyond graduation.

Career Cruising for K–12 starts by engaging elementary-aged students through a light-hearted yet informative career education program. There are two interactive game experiences, one for grades K–2 and one for grades 3–5. Students are introduced to many careers through animated characters and solve career-related mysteries as they progress through a fictional town, learning as they go. The experience results in students being more informed about different careers and gaining a basic understanding of academic and career planning, which sets the stage for them to transition to the full Career Cruising program as they move to middle and high school.

Students begin with comprehensive assessments to reveal their own strengths and interests. Then, from the database of hundreds of careers, a select list of careers is presented to the student matching his/her strengths and interests based on his/her answers. Students explore those career possibilities through multiple pathways, including interviews, job descriptions, salary expectations by region, and more. A data library gives students access to college and technical school details, including applications, tuition, enrollment, and more. Students can see which courses will best prepare them for their desired path. Customized for specific graduation requirements, course planning means that students pick their classes in alignment with their career and postsecondary goals. Students can record their goals, save preferred jobs, and track college applications in a digital portfolio. Students make their choices by test-driving career options to see how choices they make might impact their future. Financial literacy lessons help them understand the importance of financial planning and role-playing activities help pull together everything they learn.
6c. Additional Career Interest Inventories and Planning Tools

There are several other career interest inventories and career planning tools on the market that can and are being used by schools. A few examples include Kuder,\textsuperscript{75} Naviance,\textsuperscript{76} World of Work Inventory,\textsuperscript{77} and ACT Engage.\textsuperscript{78} Increasingly, career interest exploration begins as early as kindergarten and is typically provided in an online environment (see World of Work and Career Cruising).

Conclusions

Most states have adopted a definition of college and career readiness and require schools to test students using measures of readiness as part of their statewide testing programs. Definitions of readiness are typically encompassing and do not differentiate between college and career, although there is commonly a focus on academic preparation for college. Typically, states measure ELA and mathematics achievement and use those scores as the main indicators of readiness.

When states do address career paths specifically, they typically rely on WorkKeys® or narrowly defined industry credentials. Other aspects of career or work readiness, when they are addressed at all, are measured using participation rates or information about access at the school level. There is very little information generated at the student level specific to career readiness in most states.

States are increasingly demonstrating that they value social and emotional learning, but there are no widely used large-scale measures of non-cognitive skills. According to a panel of state education leaders who met at the Council of Chief State School Officers’ National Conference on Student Assessment in 2018, these skills are viewed as essential for both college and career. Skills in communicating, collaborating, creative problem solving (individually or on teams), among others, were listed as essential tools for success in a multitude of careers. The panel also indicated that these skills are very difficult to define and measure.

States are making progress towards addressing college and career readiness, but to date, their efforts have been narrow and focused on academic preparation for college, and primarily on ELA and mathematics. States tend to treat career readiness as synonymous with college readiness. When states include career readiness in their accountability systems, they tend to use narrow indicators, or they only address non-cognitive skills at the school level. The construct of readiness encompasses more than academics, and despite some progress by the states, none have fully addressed college and career readiness in their assessments and accountability systems, and information about students’ readiness remains incomplete.

\textsuperscript{75} See https://www.kuder.com/
\textsuperscript{76} See https://www.naviance.com/
\textsuperscript{77} See http://www.wowi.com/
\textsuperscript{78} See http://www.act.org/content/act/en/products-and-services/act-engage-students-parents.html
References


GOVERNING BOARD AND NAEP RESOURCES

TABLE OF RESOURCES AND LINKS

Attached documents are listed with page numbers. Click underlined links to access unattached documents online.

National Assessment Governing Board: Authority and Organization

- NAEP Law
- Board By-laws
- Board Composition and Responsibilities
- Board Members and Categories
- Ethics Primer for the National Assessment Governing Board
- Board Current Contracts
- Board Strategic Vision

NAEP Schedule of Assessments

- NAEP Schedule of Assessments
- History of Changes to the NAEP Schedule of Assessments
- News Releases

General Web-based Resources

- Home page of Governing Board web site
- Home page of the Nation's Report Card web site
- Materials for previous Board meetings

Board Policies for NAEP

- General Policy: Conducting and Reporting NAEP
- Framework Development
- Item Development and Review
- Developing Student Performance Levels for NAEP
- Reporting, Release, and Dissemination of NAEP Results
  - Guidelines for the Initial Release of The Nation's Report Card
  - Resolution on Reporting 12th Grade Academic Preparedness for College
  - Resolution on Reporting on Preparedness of 12th Grade Students
- Background Questions and the Use of Contextual Data in NAEP
- NAEP Testing and Reporting on Students with Disabilities and English Language Learners
- Trial Urban District Assessment: Eligibility Criteria and Selection Procedures
  - List of Eligible TUDA Districts
• Resolution on Linking NAEP and International Assessments

NAEP Assessment Design

• Overview of NAEP Assessment Design ................................................................. 25
• NAEP Alliance Contractors ................................................................................... 28

Selected Board Documents and Board-commissioned Research Reports (from most to least recent)

• Board Response to the National Academies of Sciences, Engineering, and Medicine 2016 Evaluation of NAEP Achievement Levels ................................................................................................................................................. 29
• Technical Report: NAEP 12th Grade Preparedness Research
• Technical Panel on 12th Grade Preparedness Research – Final Report
• The Future of 12th Grade NAEP: Report of the Ad Hoc Committee on Planning for NAEP 12th Grade Assessments in 2009
• Redesigning the National Assessment of Educational Progress

Previous “Inside NAEP” presentations

• Developing NAEP Frameworks: A Look Inside the Process
• Developing NAEP Test Questions
• Introduction to Validity
• NAEP Achievement Levels
• Sampling Concepts

Glossary of Acronyms and Other Terms .................................................................................. 35
National Assessment Governing Board

Composition
The Board is non-partisan, with 26 members representing gender, geographic, and racial-ethnic diversity. Specific categories of members specified in the NAEP law:

- **Policymakers**: governors or former governors (2), state legislators (2), chief state school officers (2), local school district superintendent (1), state (1) and local (1) school board members, nonpublic school administrator or policymaker (1)
- **Educators**: classroom teachers (3), principals (2), curriculum specialists (2)
- **Public**: general public representatives (2), parents (2), business representative (1)
- **Technical experts**: testing and measurement experts (3)

*The director of the Institute of Education Sciences serves as an ex-officio 26th member.*

Responsibilities
The responsibilities of the Board are mandated by Congress, and include:

- **Test Development**
  - Select subject areas to assess
  - Develop assessment objectives and test specifications
  - Ensure all items are free from bias
  - Have final authority on appropriateness of all items

- **Technical Methodology**
  - Develop appropriate student achievement levels
  - Design the methodology of the assessment to ensure that assessment items are valid and reliable

- **Reporting and Dissemination**
  - Develop guidelines for reporting and disseminating results
  - Plan and execute the initial public release of NAEP reports
  - Take appropriate actions needed to improve the form, content, use, and reporting of results
# National Assessment Governing Board

## Members and Categories by Term Expiration Date

<table>
<thead>
<tr>
<th>2019</th>
<th>2020</th>
<th>2021</th>
<th>2022</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alberto Carvalho</td>
<td>Rebecca Gagnon*</td>
<td>Dana Boyd</td>
<td>Paul Gasparini</td>
</tr>
<tr>
<td>Local School Superintendent</td>
<td>Local School Board Member</td>
<td>Elementary School Principal</td>
<td>Secondary School Principal</td>
</tr>
<tr>
<td>Carol Jago</td>
<td>Andrew Ho*</td>
<td>Gregory Cizek*</td>
<td>Julia Keleher</td>
</tr>
<tr>
<td>Curriculum Specialist</td>
<td>Testing &amp; Measurement Expert</td>
<td>Testing &amp; Measurement Expert</td>
<td>Chief State School Officer</td>
</tr>
<tr>
<td>Dale Nowlin*</td>
<td>Terry Mazany*</td>
<td>Tyler Cramer</td>
<td>Tonya Matthews*</td>
</tr>
<tr>
<td>Twelfth Grade Teacher</td>
<td>General Public Representative</td>
<td>General Public Representative</td>
<td>General Public Representative</td>
</tr>
<tr>
<td>Fielding Rolston*</td>
<td>Jeanette Nuñez</td>
<td>John Engler*</td>
<td>Mark Miller</td>
</tr>
<tr>
<td>State School Board Member</td>
<td>State Legislator (Republican)</td>
<td>General Public Representative</td>
<td>Eighth Grade Teacher</td>
</tr>
<tr>
<td>Linda Rosen</td>
<td>Joseph O’Keefe*</td>
<td>James Geringer*</td>
<td>Nardi Routten</td>
</tr>
<tr>
<td>Business Representative</td>
<td>Non-public School Administrator or Policymaker</td>
<td>Governor (Republican)</td>
<td>Fourth Grade Teacher</td>
</tr>
<tr>
<td>Cary Sneider*</td>
<td>Alice Peisch</td>
<td>Beverly Perdue</td>
<td></td>
</tr>
<tr>
<td>Curriculum Specialist</td>
<td>State Legislator (Democrat)</td>
<td>Governor (Democrat)</td>
<td></td>
</tr>
<tr>
<td>Ken Wagner</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chief State School Officer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joe Willhoft</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Testing &amp; Measurement Expert</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Member currently serving 2nd term and/or not eligible for reappointment.  

Updated 10/25/18
ETHICS PRIMER

FOR

THE NATIONAL ASSESSMENT
GOVERNING BOARD

November 2009
Ethics Division
Office of the General Counsel
U.S. Department of Education
EXECUTIVE SUMMARY

Now that you are a member of the National Assessment Governing Board (“NAGB”) you need to know what ethics laws and rules apply to you. The following is a very brief summary of these rules. For a more detailed discussion of how these rules apply to you, please refer to the attached summary entitled “Ethics Laws and Rules Applicable to SGEs.”

Your Status as a Special Government Employee

You are considered an SGE and not a regular federal employee because NAGB anticipates that you will be serving the federal government through your position for only 130 days or less during any period of 365 consecutive days. Whether or not you are paid by the Board for your service is irrelevant. This summary discusses how the ethics rules apply to SGEs.

Criminal Statutes Apply to Your Activities

Some of the ethics laws that apply to you carry criminal penalties. Below is a brief summary of the most important of these laws.

- The chief conflict of interest law bars you from participating personally and substantially in your capacity as a member of NAGB in any particular matter before the federal government that has a direct and predictable effect on your own financial interests or the financial interests of others with whom you have certain relationships. See 18 U.S.C. Section 208.

- If you find yourself with a financial conflict of interest, you have four options: (1) disqualify yourself (you don’t participate in any way in the matter); (2) resign from the outside entity that is the basis for the conflict; (3) sell or divest the stock or other financial interest that is the basis for the conflict; or (4) request and obtain a statutory waiver.1

- Two other laws prohibit you from representing a third party, with or without compensation, before any court or agency in connection with any particular matter involving specific parties in which the United States is a party or has a direct and substantial interest and in which you have participated personally and substantially as an SGE. In addition, if you serve the federal government for more than 60 days during the immediately preceding period of 365 consecutive days, these restrictions apply to any matter that is pending with NAGB. But remember that these restrictions do not apply to particular matters of general applicability, such as broadly applicable policies, rulemaking proceedings or legislation, that do not involve specific parties. See 18 U.S.C. Sections 203 and 205.

1 In rare circumstances, with the concurrence of the U.S. Office of Government Ethics, you may obtain a waiver of the conflict of interest.
• Another criminal law limits some of your activities after your service on NAGB ends. This law prohibits you from representing others in connection with the same particular matter involving specific parties in which you participated personally and substantially during your service to NAGB. This prohibition lasts for your lifetime. See 18 U.S.C. Section 207.

**Standards of Ethical Conduct for Employees of the Executive Branch**

The Standards of Ethical Conduct for Employees of the Executive Branch (Standards), 5 C.F.R. Part 2635, are regulations that apply both to regular federal government employees and to SGEs. However, a few exceptions exist in the Standards in recognition of the fact that SGEs are working for the government only in a very limited way. A brief synopsis of some these rules and their exceptions follow.

• **Fundraising:** You may not use your official title, position and authority to engage in fundraising.

• **Gifts:** You may not accept gifts from a “prohibited source” or offered to you because of your official position on NAGB. A prohibited source includes any person: seeking official action from NAGB; doing or seeking to do business with NAGB; conducting activities regulated by NAGB; or having interests that may be substantially affected by your official duties. There are many exceptions to this rule that are discussed in more detail in the accompanying memorandum.

• **Lobbying:** In your role as a member of NAGB, you may not urge others to contact Congress or a state legislature to urge the passage or defeat of legislation. Additional restrictions exist regarding lobbying. You should contact Department of Education’s Ethics Division before engaging in any type of lobbying.

• **Misuse of Position:** You may not use your position on NAGB or nonpublic information gained through your service on NAGB to seek advantage for yourself or others. In addition, you may not use your NAGB title in a manner that makes it appear that NAGB is sanctioning your views, products, services or personal enterprises.

• **Political Activities:** You may not engage in political activity when you are on duty or in a federal government building or car, and you may never use your official title as a member of NAGB in connection with political activities.

• **Teaching, Speaking and Writing:** You may not receive compensation for teaching, speaking or writing if: (1) the invitation was offered to you because of your position on NAGB; (2) the information conveyed by you draws substantially on nonpublic information that you obtained by working on NAGB; (3) the invitation was extended to you by an organization or person who has interests that may be substantially affected by your performance on NAGB; or (4) the subject of your work deals in a significant way
with a matter involving specific parties that you worked on while on NAGB. Again, there are some exceptions to this rule that are outlined in more detail in the accompanying memorandum.

Required Filing of a Financial Disclosure Report By SGEs

As a member of the NAGB, you are required to file a confidential financial disclosure report (also referred to as a “450” Report) when you are first appointed, and annually thereafter if you are reappointed. The purpose of the financial disclosure form is to protect you from inadvertently violating any of the criminal conflict of interest statutes and so that NAGB can know that your advice is free from any real or perceived conflicts of interest.

Please do not rely solely on this “Executive Summary” before undertaking your duties. There are many subtle nuances that are not discussed in this summary that may apply to your specific situation. The attached expanded summary provides additional detail that will help you better understand the ethics rules. Please feel free to call or e-mail Marcella Goodridge in the Ethics Division of the Office of the General Counsel at the U.S. Department of Education at (202) 401-8309, or Marcella.Keiller@ed.gov, for answers to any specific ethics questions that may arise in the course of your service on NAGB.
ETHICS LAWS AND RULES APPLICABLE TO SGES

I. INTRODUCTION

Although the ethics rules are numerous and detailed, a single, simple principle underlies these rules: *You should never use your public office for private gain, either for yourself, or for any third party.* In addition, you must refrain not only from engaging in any activity that violates the ethics rules, but you must also refrain from any activity that creates the appearance of a violation of any of these rules. The summary below is designed to help you avoid violating any ethics rules covering your activities as a member of NAGB.

II. YOUR STATUS AS A SPECIAL GOVERNMENT EMPLOYEE

A. What is a “special Government employee”?

Because you have been appointed to be a member of the NAGB and you are expected to perform your duties for not more than 130 days during the 365 days subsequent to the date of your appointment, you are, by law, a “special Government employee” (SGE). As an SGE, you are a federal government employee. This means that upon appointment, you assume the responsibilities, obligations, and restrictions that are part of public service. Because SGEs are not full-time employees, several of these restrictions apply only in limited circumstances.

B. Do the ethics restrictions apply when I am not working for NAGB?

Yes, any restrictions concerning your private activities (representational services, expert witness activities, etc.) apply equally on days when you serve the federal government through your position on NAGB and on days when you do not, except with respect to political activity. If you have not provided any services for the federal government for some time, but have not received a termination date for your appointment, you must seek a formal resolution of the matter before engaging in conduct prohibited by the ethics rules.

III. CONFLICTS OF INTEREST

A. What criminal conflict of interest statutes apply to SGES?

While you are employed as an SGE, you need to pay particular attention to four criminal conflict of interest laws found in Chapter 11, Title 18 of the United States Code: 18 U.S.C. Sections 203, 205, 207 and 208. These criminal laws include some special provisions for the treatment of SGEs. A discussion of these laws and certain related requirements found in other laws and regulations follows.
B. What financial conflicts of interest may arise for SGEs under section 208?

Section 208 prohibits you from participating personally and substantially in any particular matter that has a direct and predictable effect on your financial interests, including certain interests of others that are imputed to you under the statute. This means that you may not work on NAGB matters if you have certain connections – through the ownership of stock, through employment, or by virtue of other circumstances – with an organization that has a financial interest in the matter. For example, you may not work at all on a contract competition if you own stock valued at a certain amount in a company competing for the contract. You may not participate in a discussion of whether to modify an existing contract with a company if you work for that company. And, you may not assist in the development of a scope of work for a contract competition if you know that an organization on which you serve on the Board of Directors plans to compete for that contract.

In addition to your own personal financial interests, the financial interests of the following persons or organizations are imputed to you and also disqualify you from participating in a particular matter:

1. your spouse;
2. your minor child;
3. your general partner;
4. an organization for which you serve as an officer, director, trustee, general partner or employee; and
5. any prospective employer.

Example 1 You are on the governing board of ABC, a nonprofit organization. ABC’s financial interests are imputed to you under the statute. This means that for the purpose of determining whether you have a conflict of interest, ABC’s financial interests are treated as if they were your own. Accordingly, you may not participate in any NAGB matter in which ABC has a financial interest. Similarly, if you were in the process of discussing employment with ABC, you would be barred from participating in any NAGB matter affecting the financial interests of ABC.

Example 2 You are on the governing board of ABC (or employed by ABC, own stock in ABC, seeking employment with ABC, etc). You are asked to participate in the process of reviewing and scoring contract proposals for a contract competition for a NAGB project. Fifteen organizations have submitted a bid. When you open the proposal from one organization, you note that ABC’s name is one of the organizations that has submitted a bid. Or, perhaps ABC is listed as a subcontractor in one of the proposals. This contract competition is a “particular
matter” that will have a “direct and predictable effect” upon the financial interests of ABC. In other words, as a result of the contract competition, ABC will either gain business or not, and this decision will affect ABC financially – either negatively or positively. The amount of financial interest is not relevant – as long as ABC’s finances will be affected, unless a regulatory exemption or waiver permits you to do so, you may not work on this competition. And, because each proposal is competing against all of the others, your evaluation of competing proposals will affect the chances ABC has of winning the contract. Accordingly, you may not review any of the proposals.

You must recuse yourself from a matter as soon as you realize that you have a conflict. If, for example, you notice that you have a conflict when you are in the middle of reviewing contract proposals, you put the proposal back in its envelope and call up an NAGB staff member and let that person know that you think that you are disqualified from working on the competition. If there is any question, you should contact the U.S. Department of Education Office of the General Counsel’s Ethics Division for guidance. Once you have determined that you may not work on this matter, send the proposal back to NAGB staff.

You are permitted to participate in a particular matter affecting one campus of a multi-campus institution of higher education, where the disqualifying interest arises from your employment with a separate campus of the same institution, provided that you have no multi-campus responsibilities at the institution. If you are employed with a large university with multiple campuses and you do not have any multi-campus responsibilities, you may participate in official matters—such as grants, contracts, applications, and other particular matters—that affect the financial interests of another campus in the same university system where you are employed. Below are some examples of how section 208 may apply to your activities.

**Example 3** You are employed as a professor at the University of California-Berkeley. NAGB is planning to evaluate the impact of computer-based testing on students with disabilities and English language learners. UC-Berkeley’s science and technology department has submitted a bid. NAGB’s actions will have a direct and predictable effect on the university’s financial interest. Therefore, you may not participate in any way on this matter.

**Example 4** You are employed as a researcher at the University of California-Berkeley. NAGB is planning to evaluate the impact of computer-based testing on students with disabilities and English language learners. The University of California-Los Angeles (UCLA) has submitted a bid to be the contractor for NAGB’s evaluation. You may participate in this matter because it will not have a direct and predictable effect on either your financial interests or UC-Berkeley’s.
C. How do I resolve a conflict of interest?

1. Disqualification

A common method of resolving a conflict of interest is to disqualify yourself from participating in the matter.

*Example 5* You are serving on NAGB’s Ad Hoc Committee that will examine issues related to computer-based testing for students with disabilities and English language learners, including developing a study of computer-based testing methodologies. The Request for Proposals has been disseminated. One of the bids submitted is from ABC Corporation (ABC). You own $20,000 worth of stock in ABC. You must advise the U.S. Department of Education Office of the General Counsel’s Ethics Division that you own stock in ABC and you will not be able to participate in any way in the entire contract competition. If ABC is awarded the contract, you will also need to disqualify yourself from the entire matter.

2. Divestiture

Divestiture of a disqualifying interest (usually through the sale of stock) is another remedy available to avoid a potential violation of section 208. SGEs are not eligible for a Certificate of Divestiture (CD). A CD is a tax benefit that allows the deferral or nonrecognition of capital gain where an employee divests a financial interest in order to comply with conflict of interest requirements. Unfortunately, Congress specifically excluded SGEs from eligibility to receive CDs. 26 U.S.C. § 1043(b)(1)(A).

3. Resignation

On some very rare occasions when none of the aforementioned options are available or feasible, an SGE may need to resign from participating in an outside activity with an entity if his or her official activities as an SGE have a direct and predictable effect on the financial interest of that entity creating an irreconcilable conflict.

4. Waiver or Authorization

Another remedy to avoid a conflicting financial interest is to request and obtain a statutory waiver by contacting the Department of Education’s Ethics Division (an authorization is similar to a waiver, but only applies to non-statutory conflicts of interest - what are often referred to as “appearances of a conflict”). You may be granted a waiver only if your financial interest is not so substantial as to be deemed to be likely to affect the integrity of your services.

*Example 6* In the scenario described in Examples 1 and 2 above, you are granted a waiver permitting you to participate in a general policy matter that affects ABC’s financial interests as...
long as the matter affects all similarly situated entities in the same manner. But you would remain disqualified from participating in a matter that specifically involves ABC, which in this case means the entire contract competition.

D. What restrictions apply to my representation of third parties under sections 203 and 205?

With regard to particular matters in which you have participated personally and substantially while serving NAGB, you are prohibited from representing a third party on those particular matters, with or without compensation, before any court or agency, when the United States is a party or has a direct and substantial interest in the matter. See 18 U.S.C. Sections 203 and 205.

In addition, if you serve the federal government for more than 60 days during the immediately preceding period of 365 consecutive days, you are prohibited from representing a third party on any matter involving specific parties pending before NAGB, even if your work at NAGB did not involve these matters. These restrictions do not apply to particular matters of general applicability, such as broadly applicable policies, rulemaking procedures or legislation that does not involve specific parties.

IV. POST-EMPLOYMENT

After your appointment terminates at NAGB, you need to pay particular attention to one more criminal statute that subjects you to restrictions regarding certain matters that you may have worked on as a member of NAGB. Pursuant to 18 U.S.C. Section 207, you may never represent any third party, other than in the performance of your official government duties, in connection with the same particular matter involving specific parties in which you participated personally and substantially as a member of NAGB. This is a lifetime prohibition. For example, if you participated in a NAGB discussion concerning a contract to State University, you may never represent State University with respect to that same contract before any official of the Executive Branch of the federal government and you may never represent State University with respect to that contract in any federal court.

Further, if you serve on NAGB more than sixty days and are compensated above a certain level, you may be subject to a one-year “cooling-off” period during which you would be barred from representing before NAGB certain third parties in connection with any matter. There are some exceptions to this law as well, and you should contact the Department of Education’s Ethics Division for guidance.

V. STANDARDS OF ETHICAL CONDUCT AND OTHER ETHICS RULES

The Standards of Ethical Conduct for Employees of the Executive Branch (Standards), 5 C.F.R. Part 2635, are regulations that apply both to regular federal government employees and to SGEs. Although you are treated generally the same as regular employees under the Standards, a few
exceptions do exist for SGEs in recognition of the fact that SGEs are working for the government only in a very limited way. In addition, there are other rules that govern your conduct as an SGE, including the Hatch Act, anti-lobbying rules, the Federal Acquisition Regulation, and rules about accepting gifts and compensation from foreign governments. A brief synopsis of some of these rules follows.

A. What restrictions apply if I want to engage in fundraising?

You may not use your NAGB title, position or authority to solicit funds for any organization. In addition, you may not personally solicit funds or other support from persons whose interests may be affected substantially by the performance or nonperformance of your official duties.

B. What restrictions are there on my acceptance of gifts?

You are prohibited from accepting gifts (almost anything of monetary value) from a “prohibited source” or gifts given because of your official position as a member of NAGB, unless a specific exception applies. The definition of “prohibited source” includes any person:

- seeking official action from NAGB;
- doing or seeking to do business with NAGB; or
- having interests that may be substantially affected by your official duties at NAGB.

The definition also includes organizations the majority of whose members fall within any of these categories. You may accept various benefits resulting from your outside business or employment activities, if a reasonable person would conclude that such benefits are not offered or enhanced because of your official position. The most commonly applicable exceptions to the gift rule allow you to accept:

- Modest items of food other than a meal, such as coffee, soft drinks, or donuts;
- Most plaques, certificates and trophies;
- Discounts available to all Government employees;
- Anything for which you pay market value;
- Gifts valued at $20 or less per occasion, totaling no more than $50 in a calendar year from any one source;
- Gifts clearly motivated by friendship or family relationship;
- Gifts resulting from your outside business activities, including those of your spouse; and
- Free attendance or meal which is provided by:

1. the sponsor of the event for the day on which you are speaking at the event, or for a widely-attended gathering of mutual interest to a number of parties when the necessary determination of agency interest has been made; or

2. someone other than the sponsor of a widely-attended gathering of mutual interest to a number of parties when more than 100 people are expected to attend, the
aggregate value of the gift is under $335, and the necessary determination of agency interest has been made.

C. What restrictions apply if I want to “lobby” Congress?

NAGB and its members are permitted to communicate directly with Congress in their official capacity on matters that are related to legislation or appropriations deemed necessary to conduct NAGB’s “public business” (i.e., the NAGB’s statutory functions and responsibilities). However, the Anti-Lobbying Act, 18 U.S.C. Section 1913, prohibits you, in your official capacity at NAGB, from engaging in “grass-roots lobbying” (i.e., directly or indirectly suggesting or requesting that others contact Congress or a state legislature to urge the passage or defeat of proposed or pending legislation), even if it is related to the NAGB’s public business. The Anti-Lobbying Act also requires that any permissible direct communications with Congress in your official capacity at NAGB be made only through official channels.

None of these restrictions prohibit you from lobbying members of Congress or state legislatures, or urging others to do so, on your own time in your personal capacity. If you lobby Congress or state legislatures in your personal capacity, and the issue is related to NAGB’s business, you should make it clear that you are not representing NAGB and not acting in your official capacity as a member. Also, please note that when you are lobbying as a private citizen, you are not permitted to use government resources or equipment (including, but not limited to, computers, telephones, fax machines, copy machines, stationery), or seek assistance from NAGB staff.

D. What does “misuse of position” mean?

You may not use your position on NAGB to seek advantage for yourself or others. You also may not use nonpublic information gained through your service at NAGB to seek advantage for yourself or others. Finally, you may not use your NAGB title in a manner that makes it appear that the NAGB is sanctioning your views, products, services or personal enterprises. Of course, you may list your membership on NAGB on your curriculum vitae, but you may never use your status as an NAGB member to advertise or promote your personal activities. Please seek advice from the Department of Education Office of the General Counsel’s Ethics Division if you have any questions in this area.

E. May I keep my day job and still serve on NAGB?

Yes, you may continue to collect your regular salary from an outside employer for days on which you are providing services to the federal government (whether your federal government service is paid or unpaid). However, if you have another consultant or advisory position with NAGB or any other federal department or agency, you may not receive per diem or salary from NAGB for the same day for services performed for the two positions.

F. Are there any restrictions on my political activities?

You may not engage in any political activities while you are on duty (i.e., performing
government services) or when you are in a government building or vehicle. Although you are
not subject to any restrictions on your political activities when you are not performing
government services, you may never use your official title as a member of NAGB in connection
with any political activities.

G. What restrictions do I face if I want to teach, speak, or write on matters that are
related to the duties I perform for NAGB?

You may not receive compensation for teaching, speaking, or writing if:

- the activity is performed as part of your official duties (e.g., a speech on behalf of
  NAGB);

- the invitation to engage in the activity was extended primarily because of your
  official position at NAGB, rather than expertise in the subject matter;

- the invitation or offer of compensation was extended to you by someone with
  interests that may be affected substantially by your duties;

- the information conveyed through the activity draws substantially on nonpublic
  information obtained through your service at NAGB; or

- the activity deals, in significant part, with a matter involving specific parties to
  which you are currently assigned or had been assigned during your current
  NAGB appointment.

Notwithstanding the restrictions in bold type you may accept compensation for teaching a course
requiring multiple presentations offered as part of: (a) the regularly established curriculum of
various specified types of educational institutions; or (b) educational or training programs
sponsored and funded by federal, State, or local government. However, if you teach at an
educational institution, you must not participate in any NAGB matters that involve that
institution.

H. What restrictions apply if my government duties involve the awarding of
contracts?

If you are involved in the awarding of any contracts, please seek advice from the Ethics Division.
There are special provisions that cover your involvement in the awarding of contracts. For
example, you may not accept compensation as an employee, officer, director, or consultant of a
contractor within the one-year period after leaving Government service where you participated in
certain procurement matters pertaining to that contractor. In addition, if you disclose certain
information pertaining to Federal procurements that you obtained during your service on a
committee, you may face sanctions, including criminal penalties.
I. What restrictions apply to my interaction with foreign entities?

The emoluments clause of the U.S. Constitution prohibits you from receiving any emolument, office or title of any kind from a foreign government, including political subdivisions of a foreign government. An emolument is compensation received by virtue of holding an office or having employment with a foreign government and includes, for example, salary, honoraria, transportation, per diem allowances, household goods, shipment costs, and housing allowances. This clause has been interpreted to be broader than the traditional notion of employment and includes, for example, income received through a partnership when an identifiable portion of the partnership draw can be attributed to the partnership’s fees from such foreign government. This provision has particular relevance to positions with foreign universities that are government-operated, as opposed to private institutions. United States Constitution, art. I § 9, cl. 8. There are also statutory provisions restricting acceptance of gifts from foreign governments. 5 U.S.C. § 7342. You should seek advice from the Ethics Division regarding the details about these restrictions. Additionally, a criminal statute bars employment or consultation with a foreign entity for the purpose of providing foreign agent representation or lobbying. 18 U.S.C. § 219.

The ban on participating in foreign agent activities covered by the Foreign Agents Registration Act (FARA) prohibits representation of foreign governments or foreign political parties before the United States Government, as well as a number of other activities conducted within the United States on behalf of such entities. There are certain FARA exceptions related to trade or commerce, legal representation, humanitarian fundraising, and religious, scholastic, or scientific pursuits. The Lobbying Disclosure Act of 1995 requires certain covered Federal officials who serve as agents of foreign principals (other than foreign governments or foreign political parties) to register if they work on behalf of foreign corporations, associations, or other organizations.

Finally, certain restrictions apply after your position with NAGB terminates. Specifically, 18 U.S.C. § 207 includes restrictions on former employees who participated in trade or treaty negotiations on behalf of the United States (18 U.S.C. § 207(b)) and on former senior employees who wish to represent, or aid or advise in the representation of, a foreign entity with the intent to influence a decision of a Federal employee or agency (18 U.S.C. § 207(f)).

J. What do I do if I am called to be an expert witness?

Government employees generally may not participate as an expert witness, with or without compensation, other than on behalf of the United States, in any proceeding before a federal court or agency in which the United States is a party or has a direct and substantial interest. This restriction applies to most SGEs only if the SGE actually participated officially in the same proceeding or in the particular matter that is the subject of the proceeding. If you are appointed by the President, serve on a commission established by statute, or serve (or are expected to serve) for more than 60 days in a period of 365 days, the restriction on expert service also applies to any proceeding in which NAGB is a party or has a direct and substantial interest.
K. May I keep and use frequent flyer miles that I earn when I am on official NAGB travel?

Yes, you may use frequent flyer miles or other airline awards or promotions accumulated on official NAGB travel for your own personal use.

VI. CONCLUSION

We understand that these laws are complex and may not be intuitive. Again, we caution you that this summary is merely an introduction to the ethics laws and rules that apply to you. You should always feel free to contact the Department of Education Office of the General Counsel’s Ethics Division with any questions or concerns.

Marcella Goodridge Keiller, Attorney
U.S. Department of Education
Office of the General Counsel
400 Maryland Avenue, S.W., Room 6E237
Washington, D.C. 20202-2110
(202) 401-8309
(202) 260-5104 (fax)

Marcella.Keiller@ed.gov
### NATIONAL ASSESSMENT GOVERNING BOARD CURRENT CONTRACTS

<table>
<thead>
<tr>
<th>Contract</th>
<th>Period of Performance</th>
<th>Contractor</th>
<th>Contractor Project Director</th>
<th>Staff Member</th>
<th>Strategic Vision (SV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technical Support in Psychometrics, Assessment Development, and Preparedness for Postsecondary Endeavors</td>
<td>8/21/17 - 8/21/20</td>
<td>Human Resources Research Organization</td>
<td>Thanos Patelis</td>
<td>Sharyn Rosenberg</td>
<td>SV #2-10</td>
</tr>
<tr>
<td>Communications, Outreach, and Dissemination</td>
<td>9/29/17- 9/28/20</td>
<td>The Hatcher Group</td>
<td>Robert Johnson</td>
<td>Stephaan Harris</td>
<td>SV #1, 3, 4, &amp; 6</td>
</tr>
<tr>
<td>Developing Achievement Levels for the 2017 NAEP Grade 4 Writing Assessment</td>
<td>8/8/16 - 2/04/19</td>
<td>NCS Pearson, Inc.</td>
<td>Tim O’Neil</td>
<td>Sharyn Rosenberg</td>
<td>Legislative mandate</td>
</tr>
<tr>
<td>World Wide Web Services – <a href="http://www.nagb.gov">www.nagb.gov</a></td>
<td>6/1/15 - 1/31/19</td>
<td>Quotient, Inc.</td>
<td>Dan DeArmas</td>
<td>Stephaan Harris</td>
<td>SV #1, 3, 4, &amp; 6</td>
</tr>
<tr>
<td>Focused Reporting with NAEP Data</td>
<td>9/22/16 - 10/31/18</td>
<td>CRP, Incorporated</td>
<td>Arnold Goldstein</td>
<td>Laura LoGerfo</td>
<td>SV #1, 3, 4, &amp; 6</td>
</tr>
<tr>
<td>Review of State Mathematics Curricular Standards</td>
<td>8/16/17- 11/16/2018</td>
<td>American Institutes for Research (AIR)</td>
<td>Maria Stephens</td>
<td>Michelle Blair</td>
<td>SV #5</td>
</tr>
<tr>
<td>Update of National Assessment of Educational Progress Frameworks for Mathematics, Reading, and Other Subjects</td>
<td>9/6/2018 – 9/16/2020</td>
<td>WestEd</td>
<td>Steven Schneider</td>
<td>Michelle Blair</td>
<td>SV #5</td>
</tr>
<tr>
<td>Joint Task Force with the Council of the Great City Schools for the Trial Urban District Assessment</td>
<td>1/08/18 - 1/25/20</td>
<td>Council of the Great City Schools</td>
<td>Raymond Hart</td>
<td>Lily Clark</td>
<td>SV #1, 3, 4, &amp; 6</td>
</tr>
<tr>
<td>Joint Task Force with the Council of Chief State School Officers</td>
<td>1/26/17 - 1/25/20</td>
<td>Council of Chief State School Officers</td>
<td>Scott Norton</td>
<td>Lily Clark</td>
<td>SV #1, 3, 4, &amp; 6</td>
</tr>
<tr>
<td>Statistical Linking Studies and Related Data Sharing Agreements with Select Participating States and ACT</td>
<td>Ongoing, expected completion FY2019</td>
<td>NAEP Alliance contractors: ETS and Westat</td>
<td>NCES Liaison: Pat Etienne</td>
<td>Sharyn Rosenberg</td>
<td>SV #2</td>
</tr>
</tbody>
</table>

Updated November 2018
The Nation’s Report Card, also known as the National Assessment of Educational Progress (NAEP), was developed in 1969 to answer the important question: “How are our nation’s students doing?” The National Assessment Governing Board established this Strategic Vision to not only answer the first question, but also to expand NAEP’s impact by addressing a second question: “How can NAEP provide information about how our students are doing in the most innovative, informative, and impactful ways?”

Congress created the independent, bipartisan Governing Board in 1988 to set policy guidelines for The Nation’s Report Card, which is the largest nationally representative, continuing evaluation of the condition of education in the United States. In statute Congress charged the Governing Board to identify NAEP subjects to be tested, determine the content and achievement levels for each assessment, approve all test questions, and take steps to improve the form, reporting, and use of results.

The Governing Board partners with the National Center for Education Statistics, which administers the NAEP program, to inform a wide range of stakeholders—including policymakers, educators, researchers, business leaders, the media, and the general public—about what America’s students know and can do in various subject areas, and compare achievement data over time and among student demographic groups. This allows the nation to understand where more work must be done to improve learning among all students.

The Governing Board fulfills its statutory mission by continuously reviewing and revising its policies and practices to ensure The Nation's Report Card measures and reports meaningful information to the public.

The educational landscape of the 21st century demands increased academic ambition, greater technological sophistication, improved civic participation, and expanded global perspectives for all students. In this time of rapid and accelerating change, it is essential for The Nation’s Report Card to support innovation and address the need to improve student achievement, while maintaining its timeless promise to serve as the constant and unassailable measure of student achievement for our nation.

To increase the value of The Nation’s Report Card as a resource to impact student achievement, the Governing Board adopted this Strategic Vision with a dual focus on innovating to enhance NAEP’s form and content and informing stakeholders to expand NAEP’s dissemination and use.
The National Assessment Governing Board will promote The Nation’s Report Card’s wealth of information to facilitate the awareness and uses of NAEP in appropriate, timely, new, and meaningful ways. Examples of NAEP resources include: results; trends; test questions and tasks; studies; measurement innovations; frameworks that specify the content and design of NAEP assessments; and contextual variables about student demographics and educational experiences collected from students, teachers, and schools. The Governing Board will:

- Strengthen and expand partnerships by broadening stakeholders’ awareness of NAEP and facilitating their use of NAEP resources.
- Increase opportunities to connect NAEP to administrative data and state, national, and international student assessments.
- Expand the availability, utility, and use of NAEP resources, in part by creating new resources to inform education policy and practice.
- Promote sustained dissemination and use of NAEP information beyond Report Card releases with consideration for multiple audiences and ever-changing multi-media technologies.

The National Assessment Governing Board will revise the design, form, and content of The Nation’s Report Card using advances in technology to keep NAEP at the forefront of measuring and reporting student achievement. The Governing Board will:

- Develop new approaches to update NAEP subject area frameworks to support the Board’s responsibility to measure evolving expectations for students, while maintaining rigorous methods that support reporting student achievement trends.
- Continue improving the content, analysis, and reporting of NAEP contextual variables by considering the questions’ relevance, sensitivity, and potential to provide meaningful context and insights for policy and practice.
- Research policy and technical implications related to the future of NAEP Long-Term Trend assessments in reading and mathematics.
- Research assessments used in other countries to identify new possibilities to innovate the content, design, and reporting of NAEP.
- Develop policy approaches to revise the NAEP assessment subjects and schedule based on the nation’s evolving needs, the Board’s priorities, and NAEP funding.
- Develop new approaches to measure the complex skills required for transition to postsecondary education and career.

*This Strategic Vision will focus the work of the Governing Board through the year 2020. By pursuing these priorities, the Governing Board will ensure that The Nation’s Report Card provides the country with valuable data that measure and contribute to the improvement of student progress in achieving important knowledge and skills necessary for success as citizens in our democratic society.*

Unanimously approved November 18, 2016
The National Assessment of Educational Progress (NAEP) Authorization Act established the National Assessment Governing Board to set policy for NAEP, including determining the schedule of assessments. (P.L. 107-279)

<table>
<thead>
<tr>
<th>Year</th>
<th>Subject</th>
<th>National Grades Assessed</th>
<th>State Grades Assessed</th>
<th>TUDA Grades Assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>U.S. History*</td>
<td>8</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td>2014</td>
<td>Civics*</td>
<td>8</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td>2014</td>
<td>Geography*</td>
<td>8</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td>2014</td>
<td>TECHNOLOGY AND ENGINEERING LITERACY</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>Reading*</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td>2015</td>
<td>Mathematics*</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td>2015</td>
<td>Science**</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td>2016</td>
<td>Arts*</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td>Reading</td>
<td>4, 8</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td>2017</td>
<td>Mathematics</td>
<td>4, 8</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td>2017</td>
<td>Writing</td>
<td>4, 8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>U.S. History</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>Civics</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>Geography</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>Technology and Engineering Literacy</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2019</td>
<td>Reading</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td>2019</td>
<td>Mathematics</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td>2019</td>
<td>Science</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td>2019</td>
<td>High School Transcript Study</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2020</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2021</td>
<td>Reading</td>
<td>4, 8</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td>2021</td>
<td>Mathematics</td>
<td>4, 8</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td>2021</td>
<td>Writing</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td>2022</td>
<td>U.S. HISTORY</td>
<td>8, 12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2022</td>
<td>CIVICS</td>
<td>8, 12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2022</td>
<td>GEOGRAPHY</td>
<td>8, 12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2022</td>
<td>Economics</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2022</td>
<td>Technology and Engineering Literacy</td>
<td>8, 12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2023</td>
<td>Reading</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td>2023</td>
<td>Mathematics</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td>2023</td>
<td>Science</td>
<td>4, 8, 12</td>
<td>4, 8</td>
<td>4, 8</td>
</tr>
<tr>
<td>2023</td>
<td>High School Transcript Study</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2024</td>
<td>ARTS</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2024</td>
<td>FOREIGN LANGUAGE</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2024</td>
<td>Long-term Trend</td>
<td>(~)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NOTES:
*Assessments not administered by computer. Beginning in 2017 all operational assessments will be digitally based.
**Science in 2015 consisted of paper-and-pencil and digital-based components.
\(\sim\)Long-term Trend (LTT) assessments sample students at ages 9, 13, and 17 and are conducted in reading and mathematics.
Subjects in **BOLD ALL CAPS** indicate the year in which a new framework is implemented or assessment year for which the Governing Board will decide whether a new or updated framework is needed.
History of Changes to the NAEP Schedule of Assessments

Historical Schedule Changes
The major schedule changes adopted by the Board since 2000 are listed below:

1. Added grade 4 and 8 state-level Reading and Mathematics every two years. (2002) [Prior to the 2002 ESEA reauthorization (NCLB), state assessments at grades 4 and 8 were given every two years with reading and writing in one biennium and mathematics and science in the next, i.e., these subjects and grade 12 subjects were tested once every four years.]
2. Added the High School Transcript Study (HSTS) as a regularly scheduled study. (2005)
4. Added Technology and Engineering Literacy (TEL) to the NAEP subjects assessed. (2005)
5. Added grade 12 state-level Reading and Mathematics for volunteer states with a periodicity of every four years. (2008)
6. Adjusted the periodicity of science to correspond to the periodicity of TIMSS to conduct international benchmarking studies in mathematics and science. (2010)
7. Scheduled Writing as a technology based assessment, beginning with national data collections only and delaying fourth grade in order to complete a special study. (2010)

Other schedule changes and program adjustments from 2000 through 2015 have been due primarily to budget constraints and/or technical challenges, considering options such as:

- Assessing fewer grade levels in non-required subject areas (e.g., U.S. History, Civics, and Geography; Writing; TEL).
- Postponing a state-level assessment.
- Postponing a full assessment/study (e.g., World History, Foreign Language, HSTS).

Guiding Principles for Schedule Changes
The Governing Board’s guiding principles and priorities for schedule changes are to:

- follow the requirements in the National Assessment of Educational Progress Authorization Act, which includes the mandate to assess reading and math at the state level every two years and additional subjects as time and resources allow;
- adhere to the Governing Board’s General Policy: Conducting and Reporting the National Assessment of Educational Progress; and
- reflect the current priorities of the Governing Board to:
  - Administer all assessments using technology beginning in 2017,
  - Continue to assess broad-based curricular areas with a priority for science, technology, engineering, and mathematics (STEM); and
  - Provide state-level data in curricular areas beyond reading and mathematics.

Guidance for the schedule is found in NAEP Authorization Act Sec. 303(b)(2) which addresses the use of random sampling (A), testing in reading and mathematics at grades 4 and 8 once every two years (B), and testing in reading and mathematics at grade 12 at regularly scheduled intervals (at least as often as prior to NCLB) (C).

After this initial guidance, Sec. 303(b)(2)(D) provides guidance for including other subjects in grades 4, 8, and 12 to the extent time and resources allow. It says, including assessments “… in regularly scheduled intervals in additional subject matter, including writing, science, history, geography, civics, economics, foreign languages, and arts, and the [long term] trend assessment described in subparagraph (F).”
Overview of NAEP Assessment Design
The content and format for each NAEP subject-area assessment is determined by a NAEP assessment framework, developed under the Governing Board’s direction. General details about the structure of NAEP assessments include:

Long Test, Short Student Test Booklet
- Each student gets a small part of the test
- No individual student scores

Common Block Structures Across Subjects
- Items are within blocks, blocks are within booklets
  Example:
  At grade 4: Reading has 10 blocks and Math has 10 blocks

Test Question Types
- Multiple-choice
- Open-ended
- Computer-based tasks (Writing, Science, TEL)

Contextual Questions
- Student, teacher, administrator questionnaires

Student Booklet Block Design
While some NAEP assessments are conducted on a technology-based platform (TEL, Writing), for paper-based assessments NAEP uses a focused balanced incomplete block (BIB) or partially balanced incomplete block (pBIB) design to assign blocks or groups of cognitive items to student booklets. Because of the BIB and pBIB booklet designs and the way NAEP assigns booklets to students, NAEP can sample enough students to obtain precise results for each test question while generally consuming an average of about an hour and a half of each student's time.

The "focused" aspect of NAEP's booklet design requires that each student answer questions from only one subject area. The "BIB" or "pBIB" design ensures that students receive different interlocking sections of the assessment forms, enabling NAEP to check for any unusual interactions that may occur between different samples of students and different sets of assessment questions.

In a BIB design, the cognitive blocks are balanced; each cognitive block appears an equal number of times in every possible position. Each cognitive block is also paired with every other cognitive block in a test booklet exactly the same number of times. In a pBIB design, cognitive blocks may not appear an equal number of times in each position, or may not be paired with every other cognitive block an equal number of times. NAEP booklet design varies according to subject area (e.g., geography, mathematics, reading, science, U.S. history, writing).
Once the instrument developer has laid out the configuration of all blocks for each booklet in a booklet map shown here with the following column headings,

<table>
<thead>
<tr>
<th>Booklet number</th>
<th>Cognitive block 1</th>
<th>Cognitive block 2</th>
<th>Contextual question directions</th>
<th>General student contextual questions</th>
<th>Subject-specific contextual questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

the number of rows (booklet numbers) provides the booklet spiral design information needed for the bundling of the student booklets.


**NAEP Assessment Sample Design**

Each assessment cycle, a sample of students in designated grades within both public and private schools throughout the United States (and sometimes specified territories and possessions) is selected for assessment. In addition, in state assessment years, of which 2007 is an example, the samples of public schools and their students in each state are large enough to support state-level estimates. In all cases, the selection process utilizes a probability sample design in which every school and student has a chance to be selected, and standard errors can be calculated for the derived estimates.

**Public School Selection in State Assessment Years**

The selection of a sample of public school students for state assessment involves a complex multistage sampling design with the following stages:

- Select public schools within the designated areas,
- Select students in the relevant grades within the designated schools, and
- Allocate selected students to assessment subjects.

The Common Core of Data (CCD) file, a comprehensive list of operating public schools in each jurisdiction that is compiled each school year by the National Center for Education Statistics (NCES), is used as the sampling frame for the selection of sample schools. The CCD also contains information about grades served, enrollment, and location of each school. In addition to the CCD list, a set of specially sampled jurisdictions is contacted to determine if there are any newly formed public schools that were not included in the lists used as sampling frames. Considerable effort is expended to increase the survey coverage by locating public schools not included in the most recent CCD file.

As part of the selection process, public schools are combined into groups known as strata on the basis of various school characteristics related to achievement. These characteristics include the physical location of the school, extent of minority enrollment, state-based achievement scores, and median income of the area in which the school is located. Stratification of public schools
occurs within each state. Grouping schools within strata by such selected characteristics provides a more ordered selection process with improved reliability of the assessment results.

On average, a sample of approximately 100 grade-eligible public schools is selected within each jurisdiction; within each school, about 60 students are selected for assessment. Both of these numbers may vary somewhat, depending on the number and enrollment size of the schools in a jurisdiction, and the scope of the assessment in the particular year. Students are sampled from a roster of individual names, not by whole classrooms. The total number of schools selected is a function of the number of grades to be assessed, the number of subjects to be assessed, and the number of states participating.

Private School Selection in State Assessment Years
In years in which state-level samples are drawn for public schools, private schools are classified by type (e.g., Roman Catholic, Lutheran, etc.), and are grouped for sampling by geography (Census region), degree of urbanization of location, and minority enrollment. About 700 private schools, on average, are included, with up to 60 students per school selected for assessment. These samples are not large enough to support state-level estimates for private schools. Thus, inferences for private schools are limited to the national level, even in years when public school assessments are state-specific.

A national sample of private schools in all grades is then drawn from a list compiled through the Private School Universe Survey (PSS), which is a mail survey of all U.S. private schools carried out biennially by the U.S. Census Bureau under contract to NCES. The PSS list is updated for new schools only for a sample of Roman Catholic dioceses.

National-Only Assessment Years
In years when the NAEP samples are intended only to provide representation at the national level and not for each individual state, the public and private school selection process is somewhat different. Rather than selecting schools directly from lists of schools, the first stage of sampling involves selecting a sample of some 50 to 100 geographic primary sampling units (PSUs). Each PSU is composed of one or more counties. They vary in size considerably, and generally about 1,000 PSUs are created in total, from which a sample is selected. Within the set of selected PSUs, public and private school samples are selected using similar procedures to those described above for the direct sampling of schools from lists. The samples are clustered geographically, which results in a more efficient data collection process. The selection of PSUs is not necessary when the sample sizes are large in each state, as in state assessment years.

Source: http://nces.ed.gov/nationsreportcard/tdw/sample_design/default.aspx

NAEP Alliance Contractors
NAEP is conducted by the Assessment Division of NCES, which also works with a series of contractors. The following chart presents the structure of the collaboration between these contractors.
NAEP Alliance Contractors

To learn more about NAEP contractors in addition to the NAEP Alliance contractors, visit:
http://nces.ed.gov/nationsreportcard/contracts/history.aspx
National Assessment Governing Board’s Response to the
National Academies of Sciences, Engineering, and Medicine
2016 Evaluation of NAEP Achievement Levels

Legislative Authority

Pursuant to the National Assessment of Educational Progress (NAEP) legislation (Public Law 107-279), the National Assessment Governing Board (hereafter the Governing Board) is pleased to have this opportunity to apprise the Secretary of Education and the Congress of the Governing Board response to the recommendations of the National Academies of Sciences, Engineering, and Medicine evaluation of the NAEP achievement levels for mathematics and reading (Edley & Koenig, 2016).

The cited legislation charges the Governing Board with the authority and responsibility to “develop appropriate student achievement levels for each grade or age in each subject area to be tested.” The legislation also states that “such levels shall be determined by... a national consensus approach; used on a trial basis until the Commissioner for Education Statistics determines, as a result of an evaluation under subsection (f), that such levels are reasonable, valid, and informative to the public; ... [and] shall be updated as appropriate by the National Assessment Governing Board in consultation with the Commissioner for Education Statistics” (Public Law 107-279).

Background

NAEP is the largest nationally representative and continuing assessment of what our nation’s elementary and secondary students know and can do. Since 1969, NAEP has been the country’s foremost resource for measuring student progress and identifying differences in student achievement across student subgroups. In a time of changing state standards and assessments, NAEP serves as a trusted resource for parents, teachers, principals, policymakers, and researchers to compare student achievement across states and select large urban districts. NAEP results allow the nation to understand where more work must be done to improve learning among all students.

For 25 years, the NAEP achievement levels (Basic, Proficient, and Advanced) have been a signature feature of NAEP results. While scale scores provide information about student achievement over time and across student groups, achievement levels reflect the extent to which student performance is “good enough,” in each subject and grade, relative to aspirational goals.
Since the Governing Board began setting standards in the early 1990s, achievement levels have become a standard part of score reporting for many other assessment programs in the US and abroad.

**Governing Board Response**

**Overview**

The Governing Board appreciates the thorough, deliberative process undertaken over the past two years by the National Academies of Science, Engineering, and Medicine and the expert members of the Committee on the Evaluation of NAEP Achievement Levels for Mathematics and Reading. The Governing Board is pleased that the report concludes that the achievement levels are a meaningful and important part of NAEP reporting. The report states that, “during their 24 years [the achievement levels] have acquired meaning for NAEP’s various audiences and stakeholders; they serve as stable benchmarks for monitoring achievement trends, and they are widely used to inform public discourse and policy decisions. Users regard them as a regular, permanent feature of the NAEP reports” (Edley & Koenig, 2016; page Sum-8). The Governing Board has reviewed the seven recommendations presented in the report and finds them reasonable and thoughtful. The report will inform the Board’s future efforts to set achievement levels and communicate the meaning of NAEP Basic, Proficient, and Advanced. The recommendations intersect with two Governing Board documents, the Strategic Vision and the achievement levels policy, described here.

On November 18, 2016, the Governing Board adopted a Strategic Vision (https://www.nagb.org/content/nagb/assets/documents/newsroom/press-releases/2016/nagb-strategic-vision.pdf) to guide the work of the Board through 2020, with an emphasis on innovating to enhance NAEP’s form and content and expanding NAEP’s dissemination and use. The Strategic Vision answers the question, “How can NAEP provide information about how our students are doing in the most innovative, informative, and impactful ways?” The Governing Board is pleased that several of the report recommendations are consistent with the Board’s own vision. The Governing Board is committed to measuring the progress of our nation’s students toward their acquisition of academic knowledge, skills, and abilities relevant to this contemporary era.

The Governing Board’s approach to setting achievement levels is articulated in a policy statement, “Developing Student Performance Levels for the National Assessment of Educational Progress” (https://www.nagb.org/content/nagb/assets/documents/policies/developing-student-performance.pdf). The policy was first adopted in 1990 and was subsequently revised in 1995,
with minor wording changes made in 2007. The report motivates the revision of this policy, to add clarity and intentionality to the setting and communication of NAEP achievement levels.

The seven recommendations and the Governing Board response comprise a significant research and outreach trajectory that the Governing Board can pursue over several years in conjunction with key partners. The Governing Board will implement these responses within resource constraints and in conjunction with the priorities of the Strategic Vision.

**Evaluating the Alignment of NAEP Achievement Level Descriptors**

*Recommendation #1: Alignment among the frameworks, the item pools, the achievement-level descriptors, and the cut scores is fundamental to the validity of inferences about student achievement. In 2009, alignment was evaluated for all grades in reading and for grade 12 in mathematics, and changes were made to the achievement-level descriptors, as needed. Similar research is needed to evaluate alignment for the grade 4 and grade 8 mathematics assessments and to revise them as needed to ensure that they represent the knowledge and skills of students at each achievement level. Moreover, additional work to verify alignment for grade 4 reading and grade 12 mathematics is needed.*

The report’s primary recommendation is to evaluate the alignment, and revise if needed, the achievement level descriptors for NAEP mathematics and reading assessments in grades 4, 8, and 12. The Governing Board intends to issue a procurement for conducting studies to achieve this goal. The Governing Board has periodically conducted studies to evaluate whether the achievement level descriptors in a given subject should be revised, based on their alignment with the NAEP framework, item pool, and cut scores. The Governing Board agrees that this is a good time to ensure that current NAEP mathematics and reading achievement level descriptors align with the knowledge and skills of students in each achievement level category. In conjunction with the response to Recommendation #3, the updated Board policy on NAEP achievement levels will address the larger issue of specifying a process and timeline for conducting regular recurring reviews of the achievement level descriptions in all subjects and grades.

The Governing Board agrees strongly with the recommendation that, while evaluating alignment of achievement level descriptors is timely, it is not necessary to consider changing the cut scores or beginning a new trend line at this time. The NAEP assessments are transitioning from paper-based to digital assessments in 2017, and current efforts are focused on ensuring comparability between 2015 and 2017 scores. The Governing Board articulated this in the 2015 Resolution on Maintaining NAEP Trends with the Transition to Digital-Based Assessments ([https://www.nagb.org/content/nagb/assets/documents/policies/resolution-on-trend-and-dba.pdf](https://www.nagb.org/content/nagb/assets/documents/policies/resolution-on-trend-and-dba.pdf)).

*Recommendation #2: Once satisfactory alignment among the frameworks, the item pools, the achievement-level descriptors, and the cut scores in NAEP mathematics and reading has been
demonstrated, their designation as trial should be discontinued. This work should be completed and the results evaluated as stipulated by law: (20 U.S. Code 9622: National Assessment of Educational Progress: https://www.law.cornell.edu/uscode/text/20/9622 [September 2016]).

Ultimately, the Commissioner of Education Statistics is responsible for determining whether the “trial” designation is removed. The Governing Board is committed to providing the Commissioner with the information needed to make this determination in an expedient manner.

**Regular Recurring Reviews of the Achievement Level Descriptors**

**Recommendation #3:** To maintain the validity and usefulness of achievement levels, there should be regular recurring reviews of the achievement-level descriptors, with updates as needed, to ensure they reflect both the frameworks and the incorporation of those frameworks in NAEP assessments.

The Board’s current policy on NAEP achievement levels contains several principles and guidelines for setting achievement levels but does not address issues related to the continued use or reporting of achievement levels many years after they were established. The revised policy will seek to address this gap by including a statement of periodicity for conducting regular recurring reviews of the achievement level descriptors, with updates as needed, as called for in this recommendation. The Governing Board agrees that it is important to articulate a process and timeline for conducting regular reviews of the achievement level descriptors rather than performing such reviews on an ad hoc basis.

**Relationships Between NAEP Achievement Levels and External Measures**

**Recommendation #4:** Research is needed on the relationships between the NAEP achievement levels and concurrent or future performance on measures external to NAEP. Like the research that led to setting scale scores that represent academic preparedness for college, new research should focus on other measures of future performance, such as being on track for a college-ready high school diploma for 8th-grade students and readiness for middle school for 4th-grade students.

In addition to the extensive work that the Governing Board has conducted at grade 12 to relate NAEP mathematics and reading results to academic preparedness for college, the Governing Board has begun research at grade 8 with statistical linking studies of NAEP mathematics and reading and the ACT Explore assessments in those subjects. This work was published while the evaluation was in process and was not included in the Committee’s deliberations. Additional studies in NAEP mathematics and reading at grades 4 and 8 are beginning under contract to the National Center for Education Statistics (NCES). The Governing Board’s Strategic Vision includes an explicit goal to increase opportunities for connecting NAEP to other national and
international assessments and data. Just as the Board’s previous research related grade 12 NAEP results in mathematics and reading to students’ academic preparedness for college, the Governing Board anticipates that additional linkages with external measures will help connect the NAEP achievement levels and scale scores to other meaningful real-world indicators of current and future performance.

**Interpretations and Uses of NAEP Achievement Levels**

*Recommendation #5: Research is needed to articulate the intended interpretations and uses of the achievement levels and collect validity evidence to support these interpretations and uses. In addition, research to identify the actual interpretations and uses commonly made by NAEP’s various audiences and evaluate the validity of each of them. This information should be communicated to users with clear guidance on substantiated and unsubstantiated interpretations.*

The Governing Board’s Strategic Vision emphasizes improving the use and dissemination of NAEP results, and the Board’s work in this area will include achievement levels. The Governing Board recognizes that clarity and meaning of NAEP achievement levels (and scale scores) are of utmost importance. The Governing Board will issue a procurement to conduct research to better understand how various audiences have used and interpreted NAEP results (including achievement levels). The Governing Board will work collaboratively with NCES to provide further guidance and outreach about appropriate and inappropriate uses of NAEP achievement levels.

**Guidance for Inferences Made with Achievement Levels versus Scale Scores**

*Recommendation #6: Guidance is needed to help users determine inferences that are best made with achievement levels and those best made with scale score statistics. Such guidance should be incorporated in every report that includes achievement levels.*

The Governing Board understands that improper uses of achievement level statistics are widespread in the public domain and extend far beyond the use of NAEP data. Reports by the Governing Board and NCES have modeled appropriate use of NAEP data and will continue to do so. This recommendation is also consistent with the goal of the Strategic Vision to improve the dissemination and use of NAEP results. The Governing Board will continue to work with NCES and follow current research to provide guidance about inferences that are best made with achievement levels and those best made with scale score statistics.
**Regular Cycle for Considering Desirability of Conducting a New Standard Setting**

**Recommendation #7**: NAEP should implement a regular cycle for considering the desirability of conducting a new standard setting. Factors to consider include, but are not limited to: substantive changes in the constructs, item types, or frameworks; innovations in the modality for administering assessments; advances in standard setting methodologies; and changes in the policy environment for using NAEP results. These factors should be weighed against the downsides of interrupting the trend data and information.

When the Board’s achievement levels policy was first created and revised in the 1990s, the Board was setting standards in each subject and grade for the first time and had not yet considered the need or timeline for re-setting standards. To address this recommendation, the Governing Board will update the policy to be more explicit about conditions that require a new standard setting.

**Board’s Commitment**

The Governing Board remains committed to its congressional mandate to set “appropriate student achievement levels” for the National Assessment of Educational Progress. The Board appreciates the report’s affirmation that NAEP achievement levels have been set thoughtfully and carefully, consistent with professional guidelines for standard setting, and based on extensive technical advice from respected psychometricians and measurement specialists. The Board also takes seriously the charge to develop the current achievement levels through a national consensus approach, involving large numbers of knowledgeable teachers, curriculum specialists, business leaders, and members of the general public throughout the process. This is only fitting given the Governing Board’s own congressionally mandated membership that explicitly includes representatives from these stakeholder groups.

The Governing Board remains committed to improving the process of setting and communicating achievement levels. The Governing Board is grateful for the report recommendations that will advance these aims.

**Reference**

## Glossary of Acronyms and Other Terms

The following acronyms and terms are commonly used in the work of the National Assessment Governing Board.

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AASA</td>
<td>American Association of School Administrators</td>
</tr>
<tr>
<td>ACT</td>
<td>Formerly American College Testing</td>
</tr>
<tr>
<td>ADC</td>
<td>Assessment Development Committee <em>(Board Committee responsible for test development on all NAEP subjects)</em></td>
</tr>
<tr>
<td>AERA</td>
<td>American Educational Research Association</td>
</tr>
<tr>
<td>AFT</td>
<td>American Federation of Teachers</td>
</tr>
<tr>
<td>AIR</td>
<td>American Institutes for Research</td>
</tr>
<tr>
<td>ALDs</td>
<td>Achievement Level Descriptions</td>
</tr>
<tr>
<td>ALS</td>
<td>Achievement Levels Setting</td>
</tr>
<tr>
<td>ARRA</td>
<td>American Recovery and Reinvestment Act of 2009</td>
</tr>
<tr>
<td>AYP</td>
<td>Adequate Yearly Progress <em>(From the No Child Left Behind Act)</em></td>
</tr>
<tr>
<td>BOTA</td>
<td>Board on Testing and Assessment, National Academy of Sciences</td>
</tr>
<tr>
<td>CCSS</td>
<td>Common Core State Standards</td>
</tr>
<tr>
<td>CCSSO</td>
<td>Council of Chief State School Officers</td>
</tr>
<tr>
<td>CGCS</td>
<td>Council of the Great City Schools</td>
</tr>
<tr>
<td>COSDAM</td>
<td>Committee on Standards, Design and Methodology <em>(Board committee responsible for technical issues)</em></td>
</tr>
<tr>
<td>CRESST</td>
<td>Center for Research on Evaluation, Standards, and Student Testing <em>(Research Center at UCLA)</em></td>
</tr>
<tr>
<td>DAC</td>
<td>Design and Analysis Committee <em>(Advisory panel to ETS on technical issues in NAEP operations)</em></td>
</tr>
<tr>
<td>ECS</td>
<td>Education Commission of the States <em>(First NAEP contractor and organization supporting state policy leaders)</em></td>
</tr>
<tr>
<td>EIMAC</td>
<td>Education Information Management Advisory Consortium <em>(Advisory committee to CCSSO, mostly state testing directors)</em></td>
</tr>
<tr>
<td>ELs or ELLs</td>
<td>English Learners or English Language Learner <em>(Pronounced &quot;Ls&quot;; formerly called Limited English Proficient or LEP)</em></td>
</tr>
<tr>
<td>ELPA</td>
<td>English Language Proficiency Assessment <em>(Also ELPA21)</em></td>
</tr>
<tr>
<td>EPIC</td>
<td>Education Policy Improvement Center</td>
</tr>
<tr>
<td>ESEA</td>
<td>Elementary and Secondary Education Act</td>
</tr>
<tr>
<td>ETS</td>
<td>Educational Testing Service</td>
</tr>
<tr>
<td>FAR</td>
<td>Federal Acquisition Regulations</td>
</tr>
<tr>
<td>GAO</td>
<td>Government Accountability Office</td>
</tr>
<tr>
<td>GPO</td>
<td>Government Printing Office</td>
</tr>
<tr>
<td>GSA</td>
<td>General Services Administration</td>
</tr>
<tr>
<td>HSTS</td>
<td>High School Transcript Study <em>(A special NAEP data collection)</em></td>
</tr>
<tr>
<td>IEP</td>
<td>Individualized Education Plan <em>(A required document under the Individuals with Disabilities Education Act, which specifies learning objectives for an individual student found with a disability)</em></td>
</tr>
<tr>
<td>IES</td>
<td>Institute of Education Sciences <em>(The Department of Education office in which NCES is located. The Director of IES is an ex-officio member of the Governing Board.)</em></td>
</tr>
<tr>
<td>IRA</td>
<td>International Reading Association</td>
</tr>
<tr>
<td>-------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>IRT</td>
<td>Item Response Theory (A theory for design, analysis, and scoring of tests)</td>
</tr>
<tr>
<td>KaSA</td>
<td>Knowledge and Skills Appropriate (A series of NAEP research studies to improve measurement precision)</td>
</tr>
<tr>
<td>KSA</td>
<td>Knowledge, Skill, and/or Ability (A statement describing a subset of academic content)</td>
</tr>
<tr>
<td>LEP</td>
<td>Limited English Proficient (Term formerly used for an English Language Learner)</td>
</tr>
<tr>
<td>LTT</td>
<td>Long Term Trend Assessment (Series of NAEP tests that began in the early 1970's)</td>
</tr>
<tr>
<td>MST</td>
<td>Multi-stage Testing (A testing format where subsets of test items are presented to students based on item difficulty and student performance)</td>
</tr>
<tr>
<td>NAE</td>
<td>National Academy of Education</td>
</tr>
<tr>
<td>NAEP</td>
<td>National Assessment of Educational Progress (Pronounced &quot;nape&quot;)</td>
</tr>
<tr>
<td>NAESP</td>
<td>National Association of Elementary School Principals</td>
</tr>
<tr>
<td>NAGB</td>
<td>National Assessment Governing Board (Pronounced &quot;nag bee&quot;)</td>
</tr>
<tr>
<td>NAS</td>
<td>National Academy of Sciences</td>
</tr>
<tr>
<td>NASBE</td>
<td>National Association of State Boards of Education</td>
</tr>
<tr>
<td>NASSP</td>
<td>National Association of Secondary School Principals</td>
</tr>
<tr>
<td>The Nation’s Report Card</td>
<td>Alternate reference for NAEP assessments</td>
</tr>
<tr>
<td>NCES</td>
<td>National Center for Education Statistics (Project office for NAEP in the U.S. Department of Education and IES)</td>
</tr>
<tr>
<td>NCLB</td>
<td>No Child Left Behind Act of 2001</td>
</tr>
<tr>
<td>NCME</td>
<td>National Council on Measurement in Education</td>
</tr>
<tr>
<td>NCTE</td>
<td>National Council of Teachers of English</td>
</tr>
<tr>
<td>NCTM</td>
<td>National Council of Teachers of Mathematics</td>
</tr>
<tr>
<td>NEA</td>
<td>National Education Association</td>
</tr>
<tr>
<td>NEA</td>
<td>National Endowment for the Arts</td>
</tr>
<tr>
<td>NEH</td>
<td>National Endowment for the Humanities</td>
</tr>
<tr>
<td>NGSS</td>
<td>Next Generation Science Standards</td>
</tr>
<tr>
<td>NRC</td>
<td>National Research Council</td>
</tr>
<tr>
<td>NSBA</td>
<td>National School Boards Association</td>
</tr>
<tr>
<td>NSLP</td>
<td>National School Lunch Program</td>
</tr>
<tr>
<td>NVS</td>
<td>NAEP Validity Studies Panel</td>
</tr>
<tr>
<td>OGC</td>
<td>Office of the General Counsel (in the U.S. Department of Education)</td>
</tr>
<tr>
<td>OMB</td>
<td>Office of Management and Budget</td>
</tr>
<tr>
<td>PARCC</td>
<td>Partnership for Assessment of Readiness for College and Careers</td>
</tr>
<tr>
<td>PIRLS</td>
<td>Progress in International Reading Literacy Study</td>
</tr>
<tr>
<td>PISA</td>
<td>Program for International Student Assessment</td>
</tr>
<tr>
<td>POC</td>
<td>Principal Operating Components (Divisions of the U.S. Department of Education)</td>
</tr>
<tr>
<td>PTA</td>
<td>Parent Teacher Association</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>R&amp;D</td>
<td>Reporting and Dissemination Committee <em>(Board Committee responsible for NAEP reporting issues)</em></td>
</tr>
<tr>
<td>RFP</td>
<td>Request for Proposals</td>
</tr>
<tr>
<td>RP</td>
<td>Response probability <em>(probability of correct response on a test question)</em></td>
</tr>
<tr>
<td>RTT</td>
<td>Race to the Top <em>(also referred to as RTTT)</em></td>
</tr>
<tr>
<td>SBAC</td>
<td>SMARTER Balanced Assessment Consortium</td>
</tr>
<tr>
<td>SD</td>
<td>Students with Disabilities</td>
</tr>
<tr>
<td>SES</td>
<td>Socio-economic Status</td>
</tr>
<tr>
<td>TBA</td>
<td>Technology-based Assessment</td>
</tr>
<tr>
<td>TEL</td>
<td>Technology and Engineering Literacy <em>(A content area assessed by NAEP)</em></td>
</tr>
<tr>
<td>Department</td>
<td>United States Department of Education</td>
</tr>
<tr>
<td>Secretary</td>
<td>Secretary of Education <em>(Honorable Arne Duncan during the Obama administration)</em></td>
</tr>
<tr>
<td>TIMSS</td>
<td>Trends in International Mathematics and Science Study</td>
</tr>
<tr>
<td>TUDA</td>
<td>Trial Urban District Assessment <em>(NAEP component that measures students in large urban districts)</em></td>
</tr>
</tbody>
</table>
### NATIONAL ASSESSMENT GOVERNING BOARD

**SCHEDULE OF MEETING EVENTS: November 15-17, 2018**

Washington Court Hotel  
525 New Jersey Avenue, NW  
Washington, DC 20001  
(202) 628-2100  
Staff Office: Monticello

All meeting space is located on the Lower Lobby Level

<table>
<thead>
<tr>
<th>DATE AND TIME</th>
<th>EVENT</th>
<th>LOCATION</th>
<th>DETAILS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Thursday, November 15</strong></td>
<td>Ad Hoc Committee on Measures of Postsecondary Preparedness</td>
<td>Springwood Room</td>
<td></td>
</tr>
<tr>
<td>2:00 – 4:00 pm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Thursday, November 15</strong></td>
<td>Executive Committee</td>
<td>Ballroom 3</td>
<td>Closed Session 4:35 – 6:00 pm</td>
</tr>
<tr>
<td>4:30 – 6:00 pm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Thursday, November 15</strong></td>
<td>Informal Group Dinner</td>
<td>Joselito Casa de Comidas 660 Pennsylvania Avenue, SE Washington, DC 20003 (202) 930-6955 <a href="http://www.joselitodc.com/">http://www.joselitodc.com/</a></td>
<td>Board and staff members will walk or take local transportation to the restaurant. Attendance is optional. Please RSVP <a href="http://www.joselitodc.com/">here</a>.</td>
</tr>
<tr>
<td>6:30 – 9:00 pm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Friday, November 16</strong></td>
<td>Full Board Meeting General Session</td>
<td>Grand Ballroom</td>
<td>Committee Rooms: ADC: Montpelier COSDAM: Hermitage R&amp;D: Springwood</td>
</tr>
<tr>
<td>8:30 am – 5:15 pm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Committee Meetings</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10:30am – 12:30 pm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Working Lunch</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12:45 – 2:30 pm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Closed Session</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4:45 – 5:15 pm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Closed Session</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Friday, November 16</strong></td>
<td>Informal Group Dinner</td>
<td>701 Restaurant 701 Pennsylvania Avenue, NW Washington, DC 20004 (202) 393-0701 <a href="http://701restaurant.com/">http://701restaurant.com/</a></td>
<td>Board and staff members will walk or take local transportation to the restaurant. Attendance is optional. Please RSVP <a href="http://701restaurant.com/">here</a>.</td>
</tr>
<tr>
<td>6:00- 9:00 pm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Saturday, November 17</strong></td>
<td>Nominations Committee</td>
<td>Hermitage</td>
<td>Closed Session</td>
</tr>
<tr>
<td>7:30 – 8:20 am</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Saturday, November 17</strong></td>
<td>Full Board Meeting</td>
<td>Grand Ballroom</td>
<td>Closed Session 8:30 – 9:00 am 11:00 am – 12:00 pm</td>
</tr>
<tr>
<td>8:30 am – 12:00 pm</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Shared Ride Service
Super Shuttle provides shared ride service to and from BWI Thurgood Marshall Airport (BWI), Dulles International Airport (IAD) and Washington Reagan National Airport (DCA) to the hotel. For pick up, claim your luggage and proceed to Ground Transportation/Shared Ride Vans. Reservations are not required for transportation to the hotel. However, reservations are required for transportation to the airport. 24-hour notice is preferred, and reservations can be made on-line at www.supershuttle.com or call 1-800 BLUEVAN/ (800) 258-3826. The one-way fare is approximately $47 from BWI, $36 from Dulles and $22 from Reagan.

Taxi Service
Arrivals and Departures via BWI Thurgood Marshall and Ronald Reagan National Airports
Several taxi companies provide service from BWI Thurgood Marshall Airport (BWI) and Ronald Reagan National Airport (DCA). The one-way trip from BWI takes approximately one hour and the fare is approximately $80 - $120. The one-way fare from Reagan is approximately $20 and travel time is approximately 15 minutes. Taxi stands are located outside the airport and hotel.

Arrivals and Departures via Dulles International Airport
Washington Flyer Taxi Service (703) 661-6655 provides taxi service from Dulles International Airport. The one-way fare is approximately $60 per person and travel time is approximately 45 minutes. Upon arrival at Dulles, proceed to the baggage claim/arrivals area on the lower level of the main terminal and proceed to the Washington Flyer taxi stand. A curbside representative will assist you with coordinating service.

Public Transportation-Metrorail
The Washington Court is accessible by Metrorail via the Union Station metro station on the Red Line. Exit Union Station metro station at Massachusetts Avenue, NE & 1st Street, NW. Walk a short distance SW to 1st Street, NE. Walk south on 1st Street, NE. Turn right on Massachusetts Avenue, NE. Walk approximately 1 block and bear left on North Capitol Street, NW. Walk a short distance south on North Capitol Street, NW and turn right on F Street, NW.

Parking
Valet parking is available in the hotel's parking garage. The rates are $43 daily and $54 overnight (tax inclusive) with in and out privileges.