

## TREND IN NAEP: A PANEL DISCUSSION

The National Assessment of Educational Progress (NAEP) can measure progress only if there is stable reporting of student achievement over time. Measuring progress means maintaining trend. For each assessment administration, “maintaining trend” means that NCES can report on how students’ performance on NAEP compares with previous administrations of the same assessment. Different cohorts of students take the same assessment at the same grade as previous cohorts, but in different years. Thus, trend can be measured in how these different cohorts perform over a given timeframe.

NAEP legislation and Governing Board policy emphasize “a fair and accurate measurement of student academic achievement and reporting of trends.” Goal 1 of the Governing Board’s [\*General Policy: Conducting and Reporting The National Assessment of Educational Progress\*](#), adopted unanimously by the Board in 2013, is that NAEP must “...serve as a consistent external, independent measure of student achievement by which results across education systems can be compared at points in time and over time.”

Finally, the Board’s recently and unanimously adopted [\*Strategic Vision 2025\*](#) describes a central aspect of NAEP’s utility is “To tell the stories of American achievement for all, over time and in context.” The Strategic Vision sets two priorities to optimize this utility through changes to the NAEP assessment schedule and through updates to NAEP frameworks and assessments.

Several types of NAEP initiatives have affected the reporting of student achievement trends:

- **Inclusion**, i.e., providing accommodations to students with disabilities
- **Content**, i.e., transitioning from one framework to an updated framework
- **Format**, e.g., changing the color of text in the assessment
- **Methodology**, e.g., spiraling of assessment blocks to support matrix sampling
- **Mode**, i.e., transitioning from paper to digital assessment
- **Device**, i.e., transitioning from one digital delivery to another (e.g., laptop to tablet)

For the most part, these initiatives have resulted in trend being maintained. A key principle for maintaining trend is that the NAEP program incrementally introduces larger updates. In addition, tools used to mitigate the risk of breaking trend include:

- Using a large majority of the same items from one administration to the next
- Bridge studies (studying specific changes to NAEP, e.g., the transition from paper to digital assessment, or the transition from one content framework to another)
- Content alignment studies (to examine item pool changes from framework updates)
- Constraining content changes through Board guidance

Several international assessment programs also use these principles and tools, as well as one additional strategy: conducting content updates more frequently to avoid the accumulation of needed updates. That is, the Governing Board’s framework policy calls for framework reviews to determine the need for updates at least once every 10 years, but these reviews typically occur at

intervals longer than 10 years. However, PISA conducts framework updates every nine years, and TIMSS/PIRLS commits to updates every four years.

As the Governing Board balances priorities for framework updates and other assessment updates (e.g., in digital delivery and methodology) with Strategic Vision 2025 in mind, a key policy question is:

***How should the Board balance priorities for trend, innovation, and relevance in NAEP?***

Therefore, the November 2020 quarterly Board meeting features a panel discussion for focused discussion on this topic. Four assessment program leaders and scholars will share their perspectives. Biographies are attached.

## TREND IN NAEP: A PANEL DISCUSSION

### PANELIST BIOGRAPHIES



**Matthias von Davier** is the Monan Professor of Education at Boston College (BC) and serves as an executive director at the TIMSS & PIRLS International Study Center. Prior to joining the faculty at BC, he held the Distinguished Research Scientist position at the National Board of Medical Examiners (NBME), in Philadelphia, PA. He was a senior research director in the Research & Development Division at Educational Testing Service (ETS), and co-director of the Center for Global Assessment at ETS, leading psychometric research and operational analyses of international large scale assessments conducted by the center.

He earned his Ph.D. in psychology from University of Kiel, Germany, specializing in psychometrics. Von Davier's interests include research psychometric methodologies for analyzing data from technology-based large-scale assessments. He was one of the founding editors of the Springer journal *Large Scale Assessments in Education*. He was editor-in-chief of the *British Journal of Mathematical and Statistical Psychology (BJMSP, 2013-2018)*, and is co-editor of the Springer book series *Methodology of Educational Measurement and Assessment*. Currently, he serves as the executive editor of *Psychometrika*.



**Scott Marion** works at the National Center for the Improvement of Educational Assessment and is a national leader in conceptualizing and designing innovative and balanced assessment systems. He is also actively engaged with a broad range of Center clients including chief state school officers, legislators, state and district leaders, and classroom teachers. His projects include designing and supporting states in implementing assessment and accountability initiatives, providing technically-defensible policy guidance, and implementing high quality, locally-designed performance-based assessments. Scott coordinates and/or serves on state

or district Technical Advisory Committees (TAC) for assessment, accountability, and educator evaluation. He has served on multiple National Research Council (NRC) committees including to support designs for next generation science assessments, investigating the issues and challenges associated with incorporating value-added measures in educational accountability systems, and outlining best practices in state assessment systems. In addition, Dr. Marion serves his community as a member of the Rye (NH) School Board. He received a Ph.D. from the University of Colorado Boulder with a concentration in Measurement and Evaluation.



**Scott Norton** is the Council of Chief State School Officers' Deputy Executive Director of Programs. In this role, he oversees the programmatic work of the Council, including student expectations, student transitions, teacher workforce, and school leadership/school improvement. Scott first joined CCSSO in 2012 as the Strategic Initiative Director for Standards, Assessment, and Accountability. He previously worked as the Assistant Superintendent of the Office of Standards, Assessments, and Accountability at the Louisiana Department of Education, and he taught

public school in Louisiana. Scott holds a Ph.D. in Educational Administration and Supervision from Louisiana State University.



**Lorrie Shepard** is University Distinguished Professor in the School of Education at the University of Colorado Boulder. Her research focuses on psychometrics and the use and misuse of tests in educational settings. Most cited are her contributions to validity theory, standard setting, bias detection, the effects of high-stakes accountability testing, and the integration of learning theory with classroom formative assessment. Dr. Shepard is past president of the American Educational Research Association and past president of the National Council on Measurement in Education. She was elected to the National Academy of Education and served as its president from 2005-2009. She has served on the NAEP Validity Studies Panel since 1995. Dr. Shepard has received distinguished career awards recognizing her contributions in measurement, research, and teacher education.