Final Report on the Study to Draft Achievement-Level Descriptions for Reporting Results of the 2009 National Assessment of Educational Progress in Mathematics for Grade 12

Prepared under contract to and in conjunction with the National Assessment Governing Board

Mary Pitoniak
Gloria Dion
Dave Garber

Educational Testing Service

July 23, 2010

This report has been redacted by the National Assessment Governing Board to remove panelist names (Appendix A) and secure item-level data (Appendix C)

Listening.
Learning.
Leading

Contents

Introduction

A new framework for the National Assessment of Educational Progress (NAEP) in mathematics at grade 12 has been developed and published, and the first assessment implemented under the new framework was administered in 2009. Efforts were undertaken to determine the technical feasibility of maintaining the trend line from 2005, and the results supported the decision to maintain trend. Therefore, the trend started in 2005 will go forward with the results of the new assessment in 2009 and continue with results for subsequent assessment years.

Achievement levels are the primary means of reporting student performance for NAEP. As trend is to be maintained, the Committee on Standards, Design, and Methodology (COSDAM) recommended that the achievement-level cut scores set for grade 12 mathematics in 2005 be used for reporting results for the 2009 assessment.

The need for some changes in the achievement-level descriptions has been acknowledged. Although the number of subscales and the relative emphasis of these subscales have not changed, the 2009 framework introduced some objectives testing mathematics beyond algebra II. It also eliminated some objectives from the 2005 framework.

The study described here was focused on determining the extent to which the achievement-level descriptions provided for the 2005 framework would need to be revised to reflect the 2009 framework. The descriptions were then revised at the meeting to reflect any discrepancies, circulated for public comment, and further revised for approval by the Governing Board.

Anchoring Approach

The current anchoring study used a model-based approach[1] in which individual students are grouped in a particular achievement-level interval. After individuals are assigned to an achievement level (based on their NAEP “plausible values”), data analysts then compute the probability of each student in that achievement level answering each item correctly (or, for an open-ended question, reaching a given score level). The probabilities for students across a given level are then averaged to yield the anchoring probability used in the study for that item or score level. Each item or score level thus has four probabilities: one each for below *Basic*, *Basic*, *Proficient*, and *Advanced*.

---

[1] The model-based approach is described in detail in Appendix C of Stephen Lazer, John Mazzeo, and Andrew Weiss, *Final Report on Enhanced Achievement-Level Reporting and Scale-Anchoring Activities* (2000).

Using these processes and criteria, ETS research staff analyzed items from the 2009 NAEP mathematics assessment at grade 12[2] and determined which items mapped into given achievement-level ranges. Table 1 shows the number of items (or score points on open-ended questions) that anchored in each range. Based on their anchoring probability, items were placed into one of three anchoring categories (first set of rows in Table 1) or into one of five other categories (second set of rows in Table 1).

The items that did not anchor in one of the regions defined by the three achievement-level cut scores (those in the second set of rows in Table 1) were statistically classified based on either of two other criteria: "did not discriminate" and "did not anchor". These items either did not meet the first criterion of an anchoring probability greater than .67 for a range, or did not discriminate adequately with lower levels.

An item is viewed as being sufficiently discriminating if the difference in the item's anchoring probability at the anchor level and at the lower achievement level is greater than or equal to the 40th percentile of differences for that level. In note 9 for Figure A, the discrimination criterion is described for the item presented.

One general caveat should be offered about the data in Table 1. We often discuss whether or not "items" anchor in a given range. This is an apt depiction of any item (such as a multiple-choice question) that is scored right or wrong (i.e., a dichotomously scored item). However, items with partial credit scoring may anchor in several places. For example, for an open-ended item scored with a four-point scoring guide (scored as 1, 2, 3, or 4), there are three possible dichotomizations: (a) score 1 vs. score 2 and above, (b) score 2 and below vs. score 3 and above, and (c) score 3 and below vs. score 4. In other words, an item with a four-point guide will appear to be three (dichotomous) items in the anchoring process analysis. Clearly, these three-score-level items have quite different difficulty levels. Therefore, it is very possible that, for example, the low-score-level response to an item anchors at the *Basic* level, the middle-score-level response at the *Proficient* level, and the high-score-level response at the *Advanced* level. Similarly, an item with a three-point guide will appear to be two (dichotomous) items in the anchoring process analysis. For this reason, the total number of items (called items/score levels) in any of the columns in Table 1 is greater than the number of discrete items on the assessment.

Also, because the statistical analysis used for scale-anchoring is dependent upon the analysis used in NAEP scaling, items that failed to scale were not included in the anchoring study. For the 2009 scaling, one item was dropped. In addition, for one constructed-response item the number of score levels was collapsed due to issues that arose during scaling.

---

[2] Specifically, the items analyzed were those in the 12 new blocks developed under the 2009 framework.

Table 1

Anchoring Results for 2009 Grade 12 Mathematics

		Number Properties and Operations		**Measurement and Geometry**		**Data Analysis, Statistics, and Probability**		**Algebra**		**Total**	
Category	**Description**	**Count**	**%**	**Count**	**%**	**Count**	**%**	**Count**	**%**	**Count**	**%**
Items That Anchored at Basic, Proficient, or Advanced											
Basic	Anchors at Basic	2	10%	5	6%	10	21%	7	10%	24	11%
Prof	Anchors at Proficient	8	38%	16	19%	18	38%	26	36%	68	30%
Adv	Anchors at Advanced	5	24%	37	45%	9	19%	28	38%	79	35%
Items That Did Not Anchor											
DNA-Low	Anchors Below Basic	3	14%	1	1%	2	4%	1	1%	7	3%
DNA-Basic	Did Not Anchor at Basic (Met .67 Prob. But Not Discrimination)	0	0%	0	0%	0	0%	0	0%	0	0%
DNA-Prof	Did Not Anchor at Proficient (Met .67 Prob. But Not Discrimination)	0	0%	3	4%	2	4%	2	3%	7	3%
DNA-Adv	Did Not Anchor at Advanced (Met .67 Prob. But Not Discrimination)	0	0%	1	1%	1	2%	0	0%	2	1%
DNA-High	Did Not Anchor Because Too Difficult (Did Not Meet .67 Prob. Even at Advanced)	3	14%	20	24%	6	13%	9	12%	38	17%
	All Items	21	100%	83	100%	48	100%	73	100%	225	100%

Note. Because responses to some items were scored at multiple levels, column totals may be greater than the number of items in the assessment. Detail may not sum to totals because of rounding.

Review and Description of Items by Anchor Panel

The Anchor Panel

A panel of mathematics experts met to review the results of the scale-anchoring analysis and to produce written descriptions of the knowledge and skills displayed by students within each achievement-level range. Six panelists participated in the meeting.

Of the six panelists, three are current members of the NAEP Standing Committee and three are members of the 2009 and/or 2005 NAEP Mathematics Project committees. One is the president of a national mathematics organization. There are two high school teachers and four university-level faculty members. Appendix A provides additional biographical information about the panelists and identifies the other attendees at the anchoring meeting.

Anchor Panel Activities

The scale-anchoring meeting was held February 12–15, 2010, in Atlanta. An agenda is provided in Appendix B. The meeting began with an overview of the goals of the meeting presented by Susan Loomis of the National Assessment Governing Board. There followed a general training session by Gloria Dion, from ETS, outlining the following procedures the panelists would follow in their work of describing the assessment content:

- Review items and scoring guides.
- Discuss skills demonstrated by students responding correctly or at different levels of the scoring rubric.
- Write a descriptor of performance for each item.
- Summarize student performance at each subscale for each achievement-level range.
- Sequentially evaluate their summaries of student performance in relation to two other documents:
    - NAEP policy-level definitions, and
    - 2005 grade 12 mathematics achievement-level descriptions.
- For each of these two documents, the panelists
    - provide an initial rating of the alignment between the document and the summaries of student performance,
    - discuss their ratings, and
    - enter a second rating, which is submitted without further discussion.
- Draft achievement-level descriptions.

- Review draft achievement-level descriptions.
- Finalize achievement-level descriptions.

Following the orientation session, the panelists began writing individual item descriptors. Once each item and score level had been described, they wrote summary descriptions of what students know and can do in each of the three achievement-level ranges (*Basic*, *Proficient*, and *Advanced*). The group was facilitated by Gloria Dion of the ETS NAEP mathematics test development staff.

The panelists worked from a notebook prepared by ETS staff. The notebook contained the assessment items, scoring guides, and scale-anchoring statistics. Structured to help direct the flow of work, the items were sorted first by subscale. Then within each subscale, the items were sorted by achievement-level range (from *Basic* through *Advanced*). Within each achievement-level range, the items were arranged from easiest to most difficult. In this way, as the panelists reviewed the item pool they could see a progression in what the students knew and were able to do. Items that did not anchor were listed at the end of the book, and were not sorted by subscale.

Figure A provides an example of how the scale-anchoring data were presented for each item in the panelists' notebooks. Panelists were also given two spreadsheets for each grade: the first listed all of the items by anchor level, following the order of the items as presented in the notebook (see Appendix C); and the second listed all of the items by block to allow panelists to see where items anchored for particular blocks. The spreadsheets also included other relevant classification information such as level of mathematical complexity.

Figure A. Explanation of Scale-Anchoring Statistics

1 2 3

Anchor	Scale	Mathematical Complexity
BASIC	Data Analysis Statistics and Probability	Low

4 5 6 7 8

Item Type	# Cat	Level	Statistic	Below Basic	Basic	Proficient	Advanced	Overall
SCR	2	1	**PCT**	50.8	**88.4**	97.8	99.7	77.4
			Discrim	-	**37.6**	9.4	2.0	--

9

1. This field shows the anchoring category for the item. An item anchors at an achievement level when the estimated probability of students answering the item correctly in that level (calculated using the IRT model) is .67 or greater—*and* it meets the discrimination criterion (see note 9 for a description of the discrimination criterion). There are four anchoring categories presented in the notebook:
   - Basic—These items anchor at the *Basic* level.
   - Proficient—These items anchor at the *Proficient* level.
   - Advanced—These items anchor at the *Advanced* level.
   - DNA (Did Not Anchor)—These items did not meet the anchoring criteria, i.e., the probability of students in that level answering the item correctly did not reach .67, and/or students at that level did not meet the discrimination criteria in relation to students at the next lower level.
2. This field shows the subscale for the item. There are four subscales for the 2009 grade 12 mathematics assessment: number properties and operations; measurement and geometry; data analysis, statistics, and probability; and algebra.
3. This field shows the mathematical complexity level for each item. There are three levels of mathematical complexity for the NAEP mathematics assessment: low, moderate, and high.
4. This field shows the item type:

   MC = multiple choice
   SCR = short constructed-response
   ECR = extended constructed-response

5. This field shows the number of score categories, or levels, for the item.
6. This field shows the item score level. All multiple-choice and right/wrong constructed-response items show level 1. Multilevel constructed-response items range from 1 to 4, and should be used in conjunction with the supplied scoring guide in which they correspond to levels 2 to 5. Thus the score level on the data strip, 1, corresponds to level 2 on the scoring guide, etc.
7. The percent—PCT—for dichotomous items is the estimated probability of students answering the item correctly or reaching a given score level in that achievement level, calculated using the IRT model.

   Each score category for polytomous items is anchored separately by forming a dichotomization of the lower score levels vs. the score level and above. (For example, there are three dichotomizations for a four-point polytomous item: 1 vs. 2+3+4; 1+2 vs. 3+4; and 1+2+3 vs. 4.) PCT for these score levels refer to the probability of students at the ANCHOR level who obtain a score at the SCORE level or above.

   For this item, the average probability of students classified into the *Basic* level reaching a score level of 1 is 88.4%.
8. This field gives the overall probability of students answering this item correctly.
9. DISCRIM is the discrimination, which is the difference in probability of a correct response between each achievement level and the next lower level. Discrimination is one of the criteria used for evaluating whether an item anchors at a given level. An item is sufficiently discriminating if the difference in probability of a correct response at the anchor level and the previous anchor level is greater than or equal to the 40th percentile of differences for that level.

   For this item, the discrimination at the *Basic* level is 37.6 (88.4 minus 50.8). The value of 37.6 is above the 40th percentile value of discrimination for all items at *Basic*, which is 8.94. (The 40th percentile value is different for each achievement level. The 40th percentile values were as follows: *Basic*, 8.94; *Proficient*, 19.59; and *Advanced*, 15.29.)

Panelists began by reviewing an item, its associated anchoring data, and the scoring key or the level on the scoring guide achieved by students at the particular anchor level (e.g., *Basic*). Then, after some discussion, they described the knowledge and skills demonstrated by students who answered the question correctly. In the case of constructed-response questions, the descriptors referred to the knowledge and skills demonstrated by students receiving the particular score—for example, a "partial" or "correct" response—that anchored in the achievement-level range being reviewed. Generally, different score points on constructed-response questions anchored at different achievement levels, but when more than one score point anchored at the same level, the panelists would describe the knowledge and skills associated with the higher score point. They wrote descriptions for each credited score point for constructed-response items when the constructed-response item was first encountered, and the descriptions for each score level were placed in the appropriate level and sequence. The item-level descriptors created at the meeting are not included in this report since they contain specific information about item content.

After writing the individual descriptors for items that anchored within an achievement-level range, the panelists distilled and summarized student performance in that range for each of the subscales. To accomplish this task, they reviewed the item descriptors, grouping together those that described similar skills or knowledge. Depending on the weight of the evidence, the panelists did or did not include the topic in the summary. For example, if a number of questions measuring proportional reasoning anchored at a particular level, the panelists concluded that students performing at that level could solve problems involving proportional reasoning. If, on the other hand, students had answered only one or two questions on a topic, then panelists would omit the topic when describing what students know and can do. The summary anchor descriptions developed by the panelists, which served as a basis for their evaluations and the drafting of achievement levels, are provided in Appendix D.

Comparisons and Ratings by Anchor Panel

After completing the summary anchor descriptions, each grade-level panel was asked to make a series of comparisons between the anchor descriptions and the descriptions of performance expectations (i.e., policy level definitions and 2005 achievement-level descriptions).

For each comparison, panelists were asked to indicate for each of the three achievement levels (*Basic*, *Proficient*, and *Advanced*) whether the degree of alignment was "weak," "moderate," or "strong," and to provide comments about the degree of alignment. Panelists completed their ratings individually and, after each comparison, they discussed their ratings, providing additional comments during the discussion about the areas in which they saw alignment and lack of alignment. After the discussion of each comparison, they were asked to complete a second round of ratings. A copy of the rating forms can be found in Appendix E.

Comparison to Policy Level Definitions

Panelists first compared the 2009 anchor descriptions to the policy level definitions presented in the 2005 framework (see Appendix F). The policy level definitions are set across subject areas in NAEP and describe in very general terms what students at each grade level should know and be able to do on the assessment.[3] This comparison is intended to indicate whether performance on the new assessment, as demarcated by the cut scores, is calibrated to the policy level definitions.

Table 2 provides a summary of the ratings provided. For each achievement level, the table shows the panelists' ratings for the comparison of the 2009 anchor descriptions to the policy level definitions. For the *Basic* level, two of the panelists viewed the alignment as moderate, and four as strong. For the *Proficient* level, all of the panelists viewed the alignment as moderate. For the *Advanced* level, one panelist viewed the alignment as weak, and five viewed it as moderate.

Table 2. Comparison of 2009 Anchor Descriptions to Policy Level Definitions (Number of Panelists' Ratings)

Round 1			
	Weak (1)	**Moderate (2)**	**Strong (3)**
Basic		2	4
Proficient		6	
Advanced	1	5	
Round 2			
	Weak (1)	**Moderate (2)**	**Strong (3)**
Basic		2	4
Proficient		6	
Advanced	1	5	

Note. Ratings of 1.5 were rounded up to 2, and ratings of 2.5 were rounded up to 3.

[3] National Assessment Governing Board (2008). *Mathematics Framework for the 2009 National Assessment of Educational Progress*, Washington, DC: Author.

Comparison to 2005 Achievement-Level Descriptions

Next, panelists compared the 2009 anchor descriptions to the 2005 achievement-level descriptions (see Appendix G). The achievement-level descriptions elaborate on the generic policy definitions in describing what students at each grade level should know and be able to do on the grade 12 mathematics assessment. Panelists considered whether there was evidence that students performing within an achievement-level range have knowledge and skills that are not included in the achievement-level descriptions; or, conversely, whether there was evidence that students performing within an achievement-level range lack a specific knowledge or skill factor that is included in the achievement-level descriptions. The evaluation was intended to identify the overlap and nonoverlap in knowledge and skills between the anchor descriptions for the 2009 assessment and the achievement-level descriptions for the 2005 assessment.

Table 3 presents panelists' comparisons of the 2009 anchor descriptions to the 2005 achievement-level descriptions. For the *Basic* level, three panelists rated the alignment as moderate, and three rated it as strong. For both the *Proficient* and *Advanced* levels, four panelists rated the alignment as moderate, and two rated it as strong.

Table 3. Comparison of 2009 Anchor Descriptions to 2005 Achievement-Level Descriptions (Number of Panelists' Ratings)

Round 1			
	Weak (1)	**Moderate (2)**	**Strong (3)**
Basic		3	3
Proficient		4	2
Advanced		4	2
Round 2			
	Weak (1)	**Moderate (2)**	**Strong (3)**
Basic		3	3
Proficient		4	2
Advanced		4	2

Note. Ratings of 2.5 were rounded up to 3.

Drafting of Achievement-Level Descriptions

Following the comparisons described above, the panelists developed draft achievement-level descriptions for reporting results of the 2009 assessment.

To facilitate the development of achievement-level descriptions, panelists focused on eleven broad themes—(1) computation and symbolic manipulation; (2) proportional reasoning; (3) statistical reasoning; (4) geometric transformations; (5) functions and graphs of functions; (6) multiple representations; (7) technology; (8) mathematical justification; (9) geometry and visualization; (10) coordinate geometry; and (11) problem solving—and described what skills and knowledge were expected of students at each of the NAEP mathematics achievement levels in each of these categories. It should be noted that there were no achievement-level expectations for some themes (mathematical justification at the *Basic* level, and proportional reasoning and technology at the *Advanced* level).

Panelist Evaluations

The panelists were asked about their satisfaction with the various products resulting from the meeting. Table 4 summarizes the number of panelists rating their satisfaction at each level. The evaluation forms are included in Appendix H.

Comments from the panelists indicated overall satisfaction with the process (e.g., "This seemed like an impossible task on Friday, but was unexpectedly clear and satisfying by the end of the process."). One panelist said that it was interesting, challenging, well-organized, and generated wonderful discussions.

Table 4. Panelists' Satisfaction With Products of Meeting

	Number of Panelists' Ratings				
	Very Satisfied	**Satisfied**	**Neutral**	**Dissatisfied**	**Very Dissatisfied**
Item-Level Descriptors	3	3			
Descriptor-Based Summaries	4	2			
Achievement Level Definitions	5	1			

Meeting Summary

In summary, the scale-anchoring process proceeded in four stages. First, statistical analyses were conducted to determine the items that anchored in different achievement-level ranges. Second, a panel of mathematics experts was convened. They reviewed all items that anchored in the three different ranges and wrote individual descriptions of the mathematics skills measured by those items. The panel then created summary anchor descriptions of what students in different achievement-level ranges knew and could do. Third, the panel evaluated the alignment of the summary descriptions to the policy-level definitions and the 2005 achievement-level descriptions. Fourth, the panelists drafted achievement-level descriptions.

The 2005 achievement-level descriptions (ALDs) were used as a basis for the draft 2009 ALDs. They were modified or extended as deemed appropriate by the panelists, resulting in adjustments to the descriptions until all of the panelists were comfortable with the result.

The skills and knowledge evidenced by the items anchoring at the *Basic* level were judged by the panelist to be consistent with the skills and knowledge expected at that level. At the *Proficient* and *Advanced* levels, however, there were several instances in which the panelists felt that the skills and knowledge demonstrated by item performance were more appropriate for the lower achievement level. The panelists also carefully considered instances where knowledge and skills appeared to cross achievement levels based on the anchoring summaries.

In general, the panelists felt satisfied with the outcomes of the meeting, including the achievement-level descriptions.

Finalization of Achievement-Level Descriptions

Solicitation of Comments

After the conclusion of the meeting, the National Assessment Governing Board staff posted the draft achievement-level descriptions on their web site for public comment. The notice was posted on March 8, 2010, and feedback was requested by March 24, 2010. Comments were received from 4 individuals through the general public web posting.

In addition to the posting for public comment, the Governing board solicited comments from individuals in the Mathematics field. Comments were received from 62 individuals through this targeted solicitation.

Review of Comments

Governing Board staff forwarded the comments to ETS staff, who then sent them on those who participated in the anchor meetings. In addition to a version of the comments that was identical to that received on the web site (which were grouped by respondent), ETS staff provided a Word file with all responses to each question grouped together. Columns were provided for panelists to indicate whether action should be taken on the comment, and to record notes on their thoughts about the comment.

A web meeting was then held on March 15, 2010 to discuss panelists' recommendations on the comments. Following the web meeting, several panelists undertook the task of revising the ALDs to take into account the comments that the subpanel had deemed worth pursuing. The revised ALDs were then circulated to the panel, with many rounds of revisions going back and forth among members.

The proposed final achievement-level descriptions were brought before the Committee on Standards, Design, and Methodology of the Governing Board at their May meeting, where they were approved. The final descriptions are contained in Appendix I.

Appendix B:

Meeting Agenda

NAEP Mathematics Scale Anchoring Meeting
February 12–15, 2010
Embassy Suites Atlanta–Buckhead
3285 Peachtree Road NE
Atlanta, Georgia
404–261–7733

AGENDA

Friday, February 12

8:30 A.M. Welcome and introductions — *Gloria Dion*, *ETS*

Overview of meeting goals — *Susan Loomis*, *National Assessment Governing Board*

9:00 A.M. Description of anchor item review and general training session

10:30 A.M. Break

10:45 A.M. Review items and write descriptors for *Number Properties and Operations* items

- Discuss skills demonstrated within each anchor interval
- Write descriptors of performance for each item
- For CR items, also describe score levels that anchored within other intervals

12:00 P.M. Lunch/break time *(on your own)*

1:00 P.M. Write summary for *Number Properties and Operations* items within each anchor interval

2:00 P.M. Review items and write descriptors for *Measurement and Geometry* items

- Discuss skills demonstrated within each anchor interval
- Write descriptors of performance for each item
- For CR items, also describe score levels that anchored within other intervals

3:00 P.M. Break

3:15 P.M. Complete review for *Measurement and Geometry* items

4:30 P.M. Debriefing

5:00 P.M. Adjourn

Saturday, February 13

8:30 A.M. Write summary for *Measurement and Geometry* items within each anchor interval

10:00 A.M. Review items and write descriptors for *Data Analysis, Statistics, and Probability* items

- Discuss skills demonstrated within each anchor interval
- Write descriptors of performance for each item
- For CR items, also describe score levels that anchored within other intervals

10:30 A.M. Break

10:45 A.M. Complete review and write summary for *Data Analysis, Statistics, and Probability* items within each anchor interval

12:00 P.M. Lunch/break time *(on your own)*

1:00 P.M. Review items and write descriptors for *Algebra* items

- Discuss skills demonstrated within each anchor interval
- Write descriptors of performance for each item
- For CR items, also describe score levels that anchored within other intervals

3:00 P.M. Break

3:15 P.M. Complete review for *Algebra* items

5:00 P.M. Adjourn

Sunday, February 14

8:30 A.M. Write summary for *Algebra* items within each anchor interval

10:00 A.M. Compare and evaluate summaries across content areas by achievement level

10:30 A.M. Break

10:45 A.M. Compare and evaluate summaries across content areas by achievement level

12:00 P.M. Lunch/break time *(on your own)*

1:00 P.M. Evaluate summaries in relation to Achievement-Level Policy Definitions

2:00 P.M. Evaluate summaries in relation to 2005 Achievement-Level Descriptions

3:00 P.M. Break

3:15 P.M. Begin draft Achievement-Level Descriptions for 2009 framework

5:00 P.M. Adjourn

Monday, February 15

8:30 A.M. Continue draft Achievement-Level Descriptions

10:30 A.M. Break

10:45 A.M. Continue draft Achievement-Level Descriptions

12:00 P.M. Lunch/break time *(on your own)*

1:00 P.M. **C**omplete draft Achievement-Level Descriptions (ALDs) and Review ALDs for language consistency with ALDs for other grades

3:00 P.M. Break

3:15 P.M. Debriefing/Evaluation of Process

4:00 P.M. Meeting Adjourns

Appendix D:

Summary Anchor Descriptions

Grade 12 Mathematics Anchor Summaries	
Basic	
Number Properties and Operations	• Understand scientific notation • Perform elementary arithmetic operations with decimals
Measurement and Geometry	• Recognize single transformations, symmetry • Solve problems involving area and distance
Data Analysis, Statistics, and Probability	• Determine simple probabilities from a table, including two-way table • Recognize statistical information (univariate and bivariate) from graphical representations, including scatterplots • Use proportional reasoning to generalize from a sample to a population • Recognize that information is insufficient to solve a rate problem • Correct a graphical misrepresentation
Algebra	• Evaluate linear and quadratic functions • Read or interpret graphs of linear and generic functions
Synthesis—Basic	• Direct application of simple rules, procedures, or definitions • Single step in reasoning, evaluating, problem solving • Basic understanding of proportions • Limited decision points

Grade 12 Mathematics Anchor Summaries	
Proficient	
Number Properties and Operations	• Solve multi-step word problems • Use least common multiple in context • Computations and estimates involving exponents (including fractional exponents), absolute value, order of magnitude, and ratios • Find a counterexample to a numerical assertion • Solve problems involving rates and proportions
Measurement and Geometry	• Compute two-dimensional measurement attributes in three dimensional settings • Solve problems involving rates and proportions, including scale drawings and unit conversions • Solve multi-step word problem • Use Pythagorean Theorem and its converse, in two dimensions, including triples • Applies properties of shapes inscribed in circles • Perform single transformation in the plane • Sketch a vector given direction and speed
Data Analysis, Statistics, and Probability	• Recognize setup for computation of compound probability • Explain a graphical misrepresentation • Identify information needed to solve a rate problem presented with insufficient information • Applies elementary counting techniques • Identify appropriate formula for spreadsheet computation • Determine the effect of linear transformation on the mean of a data set • Recognize and apply the fact that standard deviation is a measure of spread • Understands the effect of outliers on summary statistics • Solve problem requiring proportional reasoning • Make inferences from graphical representations • Recognize effective means of gathering and representing data • Calculate a weighted average from a frequency table

Grade 12 Mathematics Anchor Summaries	
Proficient	
Algebra	• Evaluate radical, exponential, and piecewise-defined functions • Evaluate composition of functions (not described using formal notation) • Recognize the inverse relationship between exponential and logarithmic functions • Identify exponential functions presented verbally, graphically, or in a table • Translate from verbal to symbolic representation in linear and exponential contexts • Perform single transformation (rigid motion) of a graph in the plane • Extracts information about rate of change, slope, intercepts, or intersection from a situation presented graphically, in a table, or symbolically • Analyze a situation (presented verbally) involving non-constant rate of change and recognize a graphical model of change • Solve a system of two linear equations in two unknowns • Solve for one variable in terms of the others in given formula • Find equivalent forms of algebraic (linear, quadratic, monomial to a power) expressions • Find equivalent forms of algebraic (non-transcendental) expressions
Synthesis—*Proficient*	• Multi-step reasoning, evaluating, problem solving • Proportional reasoning • Use rules, procedures, and definitions to solve problems in more complex settings – need to decide what to do • Direct application of more sophisticated functions, rules, procedures, or definitions • Extracting information from a situation • Recognize relationships and make simple connections between ideas

Grade 12 Mathematics Anchor Summaries	
Advanced	
Number Properties and Operations	• Solve problems involving percent increase and decrease • Solve problems involving successive operations • Use algebraic reasoning in arithmetic settings • Number theoretic properties of integers
Measurement and Geometry	• Solve problems requiring visualization in three dimensions • Pythagorean Theorem to find diagonal of rectangular box • Perform sequential transformations in the plane • Use right triangle trigonometry in a variety of settings • Solve problems involving special right triangles • Solve problems using coordinate geometry • Calculate and estimate area in non-routine settings • Solve multi-step problems involving rates and proportions, including scale drawings, unit conversions, and angle measures • Reason about geometric claims and proofs • Calculate the ratio of corresponding geometric measurements of similar figures for attributes in different dimensions
Data Analysis, Statistics, and Probability	• Determine compound probability • Determine conditional probability from a two-way table • Distinguish positive and negative correlation coefficients, and relative strength • Determine equation of a line that best fits data in a scatterplot • Create appropriate formula for spreadsheet computation • Distinguish between theoretical and experimental probability • Determine the effect of linear transformation on the standard deviation of a data set

Grade 12 Mathematics Anchor Summaries	
Advanced	
Algebra	• Understands notation for function composition and evaluates composition from a table of values • Evaluate and recognize symbolic expression for functions defined recursively • Simplify rational and radical expressions • Recognize the existence of a vertical asymptote given a rational function (symbolic) • Find the inverse of a linear function • Analyze a complex problem in a coordinate geometry setting; extract multiple forms of information • Extracts and synthesizes information about rate of change, slope, intercepts, and intersection from a situation presented graphically, in a table, or symbolically • Apply definition of function • Relate verbal, graphical, and symbolic descriptions of exponential and logarithmic functions • Perform sequential transformations of a graph in the plane • Reason about parameters in the equations of lines and analyze the effects of those parameters on a graph • Identify graphical solution of a compound linear inequality in one variable • Solve a quadratic equation with complex roots (radicals) • Understand the role of hypotheses, logical implications, conclusions, and relational conjunctions in algebraic arguments • Find sum of the terms of an arithmetic sequence, in context, from verbal description
Synthesis—*Advanced*	• Evaluation, analysis, synthesis in conjunction with multi-step problem solving • Reasoning and set-up proof • Successive/sequential operations and transformations • Use more sophisticated rules, procedures, and definitions to solve problems in more complex settings – need to decide what to do

Appendix E:

Panelist Rating Forms

Panelist #: ___________

2010 NAEP Mathematics Scale Anchoring Meeting
Panelist Rating Form

Policy-Level Definitions
Round 1

Please indicate below your rating for each achievement level in regard to the following statement.

> *The degree of alignment between the content of the 2010 summary anchor descriptions and the* ***Policy-Level Definitions*** *is:*

Weak	Moderate	Strong
1	2	3

Achievement Level		
Basic	***Proficient***	***Advanced***

Comments:

____

____

____

____

____

____

Panelist #: ___________

2010 NAEP Mathematics Scale Anchoring Meeting
Panelist Rating Form

Policy-Level Definitions
Round 2

Please indicate below your rating for each achievement level in regard to the following statement.

> *The degree of alignment between the content of the 2010 summary anchor descriptions and the **Policy-Level Definitions** is:*

Weak	Moderate	Strong
1	2	3

Achievement Level		
Basic	***Proficient***	***Advanced***

Comments:

____

____

____

____

____

____

Panelist #: ___________

2010 NAEP Mathematics Scale Anchoring Meeting
Panelist Rating Form

2005 Achievement-Level Descriptions
Round 1

Please indicate below your rating for each achievement level in regard to the following statement.

The degree of alignment between the content of the 2010 summary anchor descriptions and the ***2005 Achievement-Level Descriptions*** *is:*

Weak	Moderate	Strong
1	2	3

Achievement Level		
Basic	***Proficient***	***Advanced***

Comments:

____

____

____

____

____

____

Panelist #: ___________

2010 NAEP Mathematics Scale Anchoring Meeting
Panelist Rating Form

2005 Achievement-Level Descriptions
Round 2

Please indicate below your rating for each achievement level in regard to the following statement.

> *The degree of alignment between the content of the 2010 summary anchor descriptions and the **2005 Achievement-Level Descriptions** is:*

Weak	Moderate	Strong
1	2	3

Achievement Level		
Basic	***Proficient***	***Advanced***

Comments:

____

____

____

____

____

____

Appendix F:

NAEP Policy-Level Definitions

NAEP Policy-Level Definitions

Achievement Level	Policy Definition
Advanced	This level signifies superior performance.
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.

Appendix G:

2005 Grade 12 Mathematics
Achievement Level Descriptions

2005 Grade 12 Mathematics Achievement Level Descriptions	
Basic	**Twelfth-grade students performing at the *Basic* level should be able to solve mathematical problems that require the direct application of concepts and procedures in familiar situations.** Students at grade 12 should be able to perform computations with real numbers and estimate the 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 and relationships among unknown quantities. They should be able to solve problems involving linear relations using tables, graphics, or symbols, and solve linear equations involving one variable.
Proficient	**Twelfth-grade students 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 the 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 functions ($y = k/x$); solve routine and non-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.

2005 Grade 12 Mathematics Achievement Level Descriptions	
Advanced	**Twelfth-grade students performing at the *Advanced* level should demonstrate in-depth knowledge of the mathematical concepts and procedures represented in the framework.** Students should be able to 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 on measures of central tendency and variability, analyze 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 and formulate a model for a complex situation (e.g., exponential growth and decay) and make inferences or predictions using the mathematical model.

Appendix H:

Panelist Evaluation Form

NAEP 2010 Mathematics Scale Anchoring/Achievement Level Meeting

Panelist Feedback Form

Your anonymous answers to the questions below will be used to evaluate the scale anchoring process. Thank you for completing this feedback form.

---

1. How satisfied are you with the item-level descriptors written at this meeting?

    a. Very Satisfied
    b. Satisfied
    c. Neutral
    d. Dissatisfied
    e. Very Dissatisfied

2. How satisfied are you with the descriptor-based summaries for the achievement levels written at this meeting?

    a. Very Satisfied
    b. Satisfied
    c. Neutral
    d. Dissatisfied
    e. Very Dissatisfied

3. How satisfied are you with the achievement level definitions drafted at this meeting?

    a. Very Satisfied
    b. Satisfied
    c. Neutral
    d. Dissatisfied
    e. Very Dissatisfied

(over)

4. Please provide any comments you may have on the scale anchoring process.

Appendix I:

Final NAEP Grade 12 Mathematics 2009 Achievement Level Definitions

Final NAEP 2009 Grade 12 Mathematics Achievement Level Definitions

Basic

<u>Summary of Grade 12 *Basic* Level Achievement</u>: Twelfth-grade students performing at the *Basic* level should be able to solve mathematical problems that require the direct application of concepts and procedures in familiar mathematical and real-world settings.

Students performing at the *Basic* level should be able to compute, approximate, and estimate with real numbers, including common irrational numbers. They should be able to order and compare real numbers and be able to perform routine arithmetic calculations with and without a scientific calculator or spreadsheet. They should be able to use rates and proportions to solve numeric and geometric problems.

At this level, students should be able to interpret information about functions presented in various forms, including verbal, graphical, tabular, and symbolic. They should be able to evaluate polynomial functions and recognize the graphs of linear functions. Twelfth-grade students should also understand key aspects of linear functions, such as slope and intercepts.

These students should be able to extrapolate from sample results; calculate, interpret, and use measures of center; and compute simple probabilities.

Students at this level should be able to solve problems involving area and perimeter of plane figures, including regular and irregular polygons, and involving surface area and volume of solid figures. They should also be able to solve problems using the Pythagorean theorem and using scale drawings. Twelfth graders performing at the *Basic* level should be able to estimate, calculate, and compare measures, as well as to identify and compare properties of two- and three-dimensional figures. They should be able to solve routine problems using two-dimensional coordinate geometry, including calculating slope, distance, and midpoint. They should also be able to perform single translations or reflections of geometric figures in a plane.

Proficient

<u>Summary of Grade 12 *Proficient* Level Achievement</u>: Twelfth-grade students performing at the *Proficient* level should be able to recognize when particular concepts, procedures, and strategies are appropriate, and to select, integrate, and apply them to solve problems. They should also be able to test and validate geometric and algebraic conjectures using a variety of methods, including deductive reasoning and counterexamples.

Twelfth-grade students performing at the *Proficient* level should be able to compute, approximate, and estimate the values of numeric expressions using exponents (including fractional exponents), absolute value, order of magnitude, and ratios. They should be able to apply proportional reasoning, when necessary, to solve problems in nonroutine settings, and to understand the effects of changes in scale. They should be able to predict how transformations, including changes in scale, of one quantity affect related quantities.

These students should be able to write equivalent forms of algebraic expressions, including rational expressions, and use those forms to solve equations and systems of equations. They should be able to use graphing tools and to construct formulas for spreadsheets; to use function notation; and to evaluate quadratic, rational, piecewise-defined, power, and exponential functions. At this level students should be able to recognize the graphs and families of graphs of these functions and to recognize and perform transformations on the graphs of these functions. They should be able to use properties of these functions to model and solve problems in mathematical and real-world contexts, and they should understand the benefits and limits of mathematical modeling. Twelfth-graders performing at the *Proficient* level should also be able to translate between representations of functions, including verbal, graphical, tabular, and symbolic representations; to use appropriate representations to solve problems; and to use graphing tools and to construct formulas for spreadsheets.

Students performing at this level should be able to use technology to calculate summary statistics for distributions of data. They should be able to recognize and determine a method to select a simple random sample, identify a source of bias in a sample, use measures of center and spread of distributions to make decisions and predictions, describe the impact of linear transformations and outliers on measures of center, calculate combinations and permutations to solve problems, and understand the use of the normal distribution to describe real-world situations. Twelfth-grade students should be able to use theoretical probability to predict experimental outcomes involving multiple events.

These students should be able to solve problems involving right triangle trigonometry, use visualization in three dimensions, and perform successive transformations of a geometric figure in a plane. They should be able to understand the effects of transformations, including changes in scale, on corresponding measures and to apply slope, distance, and midpoint formulas to solve problems.

Advanced

<u>Summary of Grade 12 *Advanced* Level Achievement</u>: Twelfth-grade students performing at the *Advanced* level should demonstrate in-depth knowledge of and be able to reason about mathematical concepts and procedures. They should be able to integrate this knowledge to solve nonroutine and challenging problems, provide mathematical justifications for their solutions, and make generalizations and provide mathematical justifications for those generalizations. These students should reflect on their reasoning and they should understand the role of hypotheses, deductive reasoning, and conclusions in geometric proofs and algebraic arguments made by themselves and others. Students should also demonstrate this deep knowledge and level of awareness in solving problems, using appropriate mathematical language and notation.

Students at this level should be able to reason about functions as mathematical objects. They should be able to evaluate logarithmic and trigonometric functions and recognize the properties and graphs of these functions. They should be able to use properties of functions to analyze relationships and to determine and construct appropriate representations for solving problems, including the use of advanced features of graphing calculators and spreadsheets.

These students should be able to describe the impact of linear transformations and outliers on measures of spread (including standard deviation), analyze predictions based on multiple data sets, and apply probability and statistical reasoning to solve problems involving conditional probability and compound probability.

Twelfth grade students performing at the *Advanced* level should be able to solve problems and analyze properties of three-dimensional figures. They should be able to describe the effects of transformations of geometric figures in a plane or in three dimensions, to reason about geometric properties using coordinate geometry, and to do computations with vectors and to use vectors to represent magnitude and direction.