

NATIONAL ASSESSMENT GOVERNING BOARD

WORK STATEMENT

FOR

DEVELOPING ACHIEVEMENT LEVELS

ON THE 2009

NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

FOR SCIENCE

GRADES 4, 8, AND 12

National Assessment Governing Board

July 8, 2008

TABLE OF CONTENTS

1. INTRODUCTION	1
2. BACKGROUND	1
2.1 History of NAEP.....	1
2.2 Enabling Legislation for NAEP Achievement Levels.....	1
2.3 National Assessment Governing Board Policy on Achievement Levels.....	2
2.4 Achievement Levels Activities 1992 through 2008.....	3
2.5 Overview of Technical and Procedural Issues.....	4
2.6 Technical Direction of the Contract.....	6
2.7 Item Security and Data Release	6
3. SCOPE OF WORK.....	7
3.1 Overview of the Procurement	7
3.2 Procurement Tasks.....	9
3.2.1 Task 1: Develop Planning Document	9
3.2.2 Task 2: Design Achievement Levels-setting Procedures and Produce a Design Document.....	9
3.2.3 Task 3: Conduct Achievement Levels-Setting Process	12
3.2.4 Task 4: Develop Final Reports	15
3.3 Schedule of Project Deliverables	17
4. INSTRUCTIONS TO OFFERORS	19
4.1 Organization and Content of the Technical Proposal	19
4.1.1 Table of Contents.....	20
4.1.2 Introduction.....	20
4.1.3 General Approach	20
4.1.4 Technical Work Plan.....	20
4.1.5 Management Plan.....	20
4.1.6 Related Experience of Proposed Staff	22
4.1.7 Related Organizational Experience.....	22
4.2 Business Proposal	23
4.3 Use of Subcontractors	23
4.4 Period of Performance	24
APPENDICES	25
APPENDIX A: Achievement Levels Policy and Implementation Guidelines	26
APPENDIX B: Website Locations of Resource Materials.....	39
APPENDIX C: Sample Invoice	40
APPENDIX D: Schedule of Board Quarterly Meetings 2008-2010.....	42

1. INTRODUCTION

The purpose of this procurement is to implement a process to produce a set of recommendations which will assist the National Assessment Governing Board in developing achievement levels for the 2009 Science National Assessment of Educational Progress (NAEP). Specifically, the requirement is for designing and conducting an achievement levels-setting activity for the 2009 NAEP for grades 4, 8, and 12 in science in accordance with the Board's policies on standard setting, and for conducting related studies.

The achievement levels shall be established by the Board to define what students should know and be able to do in science at grades 4, 8, and 12. The achievement levels will be used for reporting NAEP results to the American public. In arriving at a policy decision on science achievement levels, the National Assessment Governing Board will be informed by the recommendations for the achievement levels produced under this contract.

2. BACKGROUND

2.1 History of NAEP

For nearly four decades, NAEP has provided information on the achievement and performance of students in the United States. The National Center for Education Statistics (NCES), within the Department of Education, has responsibility for administering NAEP. The National Assessment Governing Board (hereafter referred to as the Governing Board or the Board) is responsible for setting policy for NAEP.

For most assessments, nationally representative samples of students (from 35,000 to 100,000) are drawn from three grade levels for assessments in various subject areas. The resulting data on student knowledge and performance have been accompanied by descriptive background information allowing analyses of a variety of student experiences and background factors that relate to student achievement. NAEP, known as "The Nation's Report Card," is the most comprehensive and only continuing, valid source of information on what U.S. students know and can do. Tracking student performance has permitted the nation to assess American students' academic progress over time.

2.2 Enabling Legislation for NAEP Achievement Levels

Public Law 107-279— *The National Education Statistics Act of 1994*, as amended by the *No Child Left Behind Act of 2001*, contains the authorizing legislation for the NAEP as well as the Governing Board. (Please see [The NAEP Law](#) section under the *NAEP* tab at www.nagb.org.) The following are among the specific responsibilities assigned to the Governing Board in the legislation:

- Select the subject areas to be assessed (consistent with section 303(b));
- Develop appropriate student achievement levels as provided in section 303(e);
- Develop guidelines for reporting and disseminating results;
- Develop standards and procedures for regional and national comparisons.

The legislation specifies that the Board is responsible for 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, as well as for devising a national consensus approach. The legislation further specifies that the levels “shall be used on a trial basis ... and updated as appropriate.”

The achievement levels remain in trial status, and all recent NAEP Report Cards contain language similar to the following statement in The Nation’s Report Card for Science 2005:¹

As provided by law, NCES, upon review of congressionally mandated evaluations of NAEP, has determined that achievement levels are to be used on a trial basis and should be interpreted with caution. However, NCES and the Governing Board have affirmed the usefulness of these performance standards for understanding trends in achievement. NAEP achievement levels have been widely used by national and state officials.

2.3 National Assessment Governing Board Policy on Achievement Levels

To accomplish its legislated responsibilities, the Board has established a policy framework (amended March 4, 1995) entitled *Developing Student Performance Levels on the National Assessment of Educational Progress*. The policy (See Appendix A) specifies:

The purpose for developing student performance levels on NAEP is to clarify for all readers and users of NAEP data that these are expectations which stipulate *what students should know and should be able to do* at each grade level and in each content area measured by NAEP.

The policy specifies three achievement levels, as follows:

The levels-setting process shall produce three threshold points for each content area and at each grade level assessed, demarcating entry into three categories: *basic, proficient, and advanced*. These levels are defined as:

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.

¹ SOURCE: Grigg, W., Lauko, M., and Brockway, D. (2006). *The Nation’s Report Card: Science 2005*. (NCES 2006-466). U.S. Department of Education, National Center for Education Statistics. Washington, D.C.: U.S. Government Printing Office. p.5.

It also specifies:

Developing achievement levels shall be a widely inclusive activity of the Board, utilizing a national consensus approach, and providing for the active participation of teachers, other educators (including curriculum specialists and school administrators at the local and state levels), and noneducators including parents, members of the general public, and specialists in the particular content area;

and:

Proposed achievement levels shall be reviewed by a broad constituency, including consumers of NAEP data, such as policymakers, professional groups, the states and territories. In carrying out its responsibilities, the Board will ordinarily engage the services of a contractor who will prepare recommendations for the Board's consideration on the levels, the descriptions, and the exemplar exercises.

The Board's Achievement Level policy and the guidelines regarding implementation of the policy are provided in Appendix A.

In addition to the policy highlights and definitions presented above, technical proposals submitted in response to the Request for Proposals shall reflect the entire scope of the Board policy in discussing issues and design features of the levels-setting activity for science in 2009.

2.4 Achievement Levels Activities 1992 through 2008

The Governing Board has completed comprehensive processes to set achievement levels on the 1992 NAEP assessments in mathematics and reading, the 1994 assessments in U.S. history and geography, the 1996 assessment in science, the 1998 assessments in civics and writing, the 2005 assessment in twelfth grade mathematics, and the 2006 assessment in twelfth grade economics. These levels have been used in reporting NAEP results for the subjects in subsequent assessment years (e.g., in 1994, 1996, 2000, 2004, 2005, and 2007 Report Cards). In addition to national data, assessments in mathematics, reading, writing, and science provide state-level information on school performance based on comparable, representative samples of students at grades four and eight. The Trial Urban District Assessment (TUDA) currently includes ten large urban districts, and seven districts will be added to the assessments of grades 4 and 8 in reading, mathematics, and science in 2009. Although twelfth grade is not a part of the main NAEP state assessment program, a pilot study including eleven states is planned for the 2009 assessments of grade 12 reading and mathematics.

The achievement levels-setting procedures to date were accomplished under the direction of the Board through contracts with ACT Inc. (Formerly American College Testing). This work was accomplished with the assistance of a number of technical teams; input from a broad spectrum of educational, assessment, business, and policy groups and the general public; involvement of local, state, and federal individuals and organizations; and evaluations by various technical teams and advisory panels. The achievement levels-setting process was implemented in accordance with the policy framework, definitions, and technical procedures outlined in the Board's Achievement Levels policy. Additional technical and procedural issues arising from previous levels-setting efforts are presented in the next section.

Previous achievement levels activities have provided an excellent foundation upon which to base further levels-setting efforts. Proposals shall consider the findings of these major levels-setting activities in the design of achievement levels-setting procedures for the science assessment. Basic tenets and procedures in the Board's policy on achievement levels shall be followed throughout.

The Board is committed to reporting achievement level data for all subjects and grades tested in the National Assessment. The 2009 NAEP Science Framework and the Test Specifications, adopted by the Governing Board, are available on the Governing Board's website (www.nagb.org/frameworks). The framework and test specifications for science focus on developing a comprehensive and challenging assessment. The science content measured by NAEP was informed by international assessments, the national content standards, science content included in state standards, research on learning and assessment in science, and best practices. The science framework was widely distributed for public review and evaluation, and the comments were very positive. The science NAEP has included "hands-on" tasks for students since 1996. The new framework and test specifications call for interactive computer-based tasks as well as hands-on tasks to assess students' ability to conduct scientific research and use analytic cognitive skills. While these lab-type tasks are an innovative and exciting feature of the science NAEP, performance of students on the hands-on and computer-based tasks will *not* be included in computation of the overall performance score; they will *not* be included in the achievement levels-setting process.

2.5 Overview of Technical and Procedural Issues

During the process of setting achievement levels for NAEP, several technical and procedural issues have been raised and discussed by the Board. These issues, in addition to others identified by the offerors, should be addressed in the design for setting achievement levels on the 2009 science assessment.

▪ Standard Setting Methodology

Sound standard-setting methodologies shall be used for the 2009 science achievement levels-setting (ALS) activities. The methods employed will have a solid research base and will exhibit all the technical characteristics necessary to meet the current psychometric standards in the field.

The Governing Board has extensive experience with the modified Angoff procedure for NAEP achievement levels setting. The final version of the evolution of the methodology for NAEP is represented by the 1998 ALS procedures for civics and writing. Modifications of the bookmark process, called Mapmark, were used for setting achievement levels for the grade 12 mathematics NAEP in 2005 and for the grade 12 economics NAEP in 2006. This procurement for science achievement levels-setting work is designed to further the information available to the Governing Board regarding an item mapping procedure, such as the Bookmark process, and ways to enhance the procedural validity and internal consistency of the process.

The Mapmark procedure used for NAEP incorporates item maps to show the location of each item with respect to the score scale, given a mapping criterion of $RP=.67$.² The Board has approved the use of this mapping criterion in previous achievement levels-setting procedures and wishes to have items mapped to scale scores using this criterion in the bookmark/item mapping procedures implemented for science. Because of the importance of the choice of an RP criterion, the Board remains interested in more information regarding panelists' understanding of the RP criterion and of how the RP criterion impacts the outcomes of the achievement levels-setting process. Thus, the Board wishes to have information collected from panelists to address specifically the extent to which panelists understand the following:

1. How participants understand and use a given RP criterion and how the cognitive demand of considering an RP criterion compares to the cognitive demands of using other sources of information in the standard setting process;
2. How the cutscore would change with an increase or decrease in the RP criterion used for placing the bookmark;
3. How the knowledge, skills, and abilities represented by items that map within an achievement level score range relate to the achievement levels definitions of what students *should* know and be able to do.

The Board is interested in informing the judgments of panelists efficiently and effectively through the provision of feedback. Proposals should provide detailed information regarding the feedback to be provided to inform panelists about their judgments, and provide a rationale for the feedback information, including the timing and frequency of such feedback.

- **Composition, Size, and Timing of Panels**

The composition of the achievement levels panels—two-thirds educators and one-third noneducators — has been deemed appropriate by the Board since the beginning of its standard setting initiative in 1992. According to Board policy, the panelists must represent a balance of demographic characteristics such as gender, race/ethnicity, and geographic region. Sufficient time is essential for the panelists to receive training and perform the ALS tasks appropriately. In addition, the scheduling of meetings should provide for the maximum time possible for all components of the achievement levels process, including Board direction and decision-making, as well as the preparation of reports of the NAEP results in terms of the achievement levels.

- **Judgments Using Empirical Data**

Consistent with the Board's policy, the achievement level panelists should be presented with extensive information regarding the assessment. Participants shall be presented with actual student performance on the assessment items at an appropriate time during the levels-setting activities. The rationale for providing panelists with the specific empirical data, the timing of the introduction of such information within the process, and the impact on the panelists' ratings must be clearly specified and documented. Offerors are encouraged to recommend data to inform the panelists' judgments during the levels-setting activities and to inform the Board's decision regarding the achievement levels.

² RP stands for "response probability," and $RP .67$ means that there is a .67 probability of correct response.

- **Describing the Achievement Levels**

The Governing Board will engage content experts, outside the scope of this procurement, to develop final achievement level descriptions from the preliminary descriptions that appear in the 2009 NAEP Science Framework (www.nagb.org/frameworks). The Governing Board will provide the final recommended descriptions to the contractor to use in the ALS process. The descriptions are the fundamental component of NAEP achievement levels, and they shall be central to the judgmental process.

2.6 Technical Direction of the Contract

The process by which the National Assessment Governing Board reaches decisions on the achievement levels for the science assessments will be informed by the technical work of this contract. For achievement levels-setting initiatives, the Board has assigned the Committee on Standards, Design and Methodology (COSDAM) the responsibility to monitor and oversee the levels-setting process. The performance of work requested in this RFP shall be subject to the technical direction of the Governing Board Contracting Officer (CO), the Contracting Officer's Representative (COR), the COSDAM, and the Board. Such technical direction shall consist of, but shall not be limited to the following activities:

- Monitoring the contractor's performance to ensure compliance with the technical requirements of the contract;
- Providing direction to the contractor to ensure compliance with appropriate Board policies and specifications as contained in the appendices to this RFP and as communicated through regularly scheduled reviews of work-in-progress by COSDAM and the Board; and
- Recommending to the Contracting Officer final acceptance or rejection of all deliverables based on the review and disposition of deliverables by COSDAM and the Board.

The contractor shall work closely with the COR in all aspects of the work and communicate at least once per week to report on the project status. The ability of the contractor's staff to communicate effectively and productively with the COR is critical. The COR will closely monitor the contractor's work to ensure performance of activities as specified in the contract and to guarantee technical quality, policy relevance, and objectivity. All contract-related communications or proposed changes in the scope of work will be directed to the Contracting Officer.

2.7 Item Security and Data Release

The Governing Board is especially sensitive to the fact that many groups and individuals will have access to various parts of the NAEP item pool during the achievement levels-setting process, and they will have information on item data prior to the official release of data. Ensuring security is a serious concern, and it is the contractor's responsibility to guarantee and maintain security for NAEP materials.

It is the responsibility of the contractor to demonstrate how security procedures will be employed. The contractor will specifically address security procedures for (1) item distribution, (2) item review, (3) data review, and (4) hotel and other staff security maintenance. The contractor will include security procedures and plans in a Design Document to be produced as specified under Task 2 in the Scope of Work, described in Section 3.2.2 below. These plans must be approved in writing by the COR before implementation.

NCES requires that any person(s) who will be reviewing or using secure data or materials sign compliance forms during the course of this contract. These forms will be supplied by the COR for the contractor to complete. The contractor shall forward two copies of all compliance forms to the COR who will forward one copy to the NCES COR for the NAEP Design, Analysis, and Reporting (DAR) contract.

3. SCOPE OF WORK

3.1 Overview of the Procurement

In setting achievement levels for the 2009 science assessment and in planning for student performance standards, the contractor shall perform the following tasks and provide the following services:

(A) Design a process to assist the Board in setting achievement levels on the science assessment for grades 4, 8, and 12. The process must conform to the policy framework, definitions, and technical procedures outlined in the Board's policy on setting achievement levels. (See Appendix A.) This policy states that the Board will define three levels of achievement — basic, proficient, and advanced — for each grade and subject tested by the National Assessment.

For this procurement, the offeror is required to design a procedure for implementing a bookmark-type methodology for setting science achievement levels at three grade levels modified to comply with the specific requirements of Governing Board policies and this statement of work.

In addition, the Governing Board wishes to address panelists' understanding of specific aspects of the bookmark methodology, as noted in Section 2.5. The study design should produce an unambiguous indication of the feasibility of each type of feedback information to be provided and provide a rationale for the timing and choice of procedures.

(B) Implement a pilot study using the exact procedures designed for setting the 2009 NAEP Science achievement levels. (See Appendix B for the website location of the process report which includes information about the economics pilot study.) A smaller number of panelists may be used for the pilot study, but the number shall be sufficient to produce reliable results.

(C) Conduct the achievement levels-setting process to use the informed judgments of well-qualified and broadly representative panels to recommend achievement level cut scores consistent with the achievement levels definitions for basic, proficient, and advanced levels and to identify exemplar items for each level.

- (D) Capture and document the elements of the achievement levels-setting process and panelists' reactions to each. The extent to which this information provides evidence of procedural validity shall be evaluated and reported.
- (E) Report validity information to assist the Board in evaluating the recommendations resulting from the achievement levels-setting process and in reaching a policy decision regarding the achievement levels to use for reporting student performance on the 2009 science NAEP.
- (F) Produce a final report of the recommended descriptions, cut scores, and exemplar items for the achievement levels that documents the process and provides recommendations for future levels-setting efforts. The Governing Board plans to distribute these reports widely to various assessment and subject-matter stakeholders. Thus, the reports must be technically precise and complete, as well as comprehensible to a wide-ranging audience. The report quality should meet professional standards for publication.
- (G) Ensure the effectiveness of the achievement levels-setting process. The contractor, under the technical direction of the COR and with oversight by the Committee on Standards, Design and Methodology, shall:
1. Address the policy and technical issues relevant to establishing achievement levels for the 2009 NAEP Science in the proposal.
 2. Design achievement levels-setting activities to conform to the Board's policies on achievement levels, with appropriate modifications for the unique design and format of the science assessment and other requirements specified in the procurement.
 3. Obtain comment from a broad spectrum of educators, policymakers, and the general public on the levels-setting design described in the Design Document specified in Task 2 of the Scope of Work.
 4. Engage a broadly representative and well-qualified group of state and local educators (public and private), scholars, scientists, employers, business and civic group representatives, and other interested citizens to serve as panelists in carrying out the achievement levels-setting activities.
 5. Conduct the achievement levels-setting activities; perform analyses on the data obtained from the procedures; produce achievement levels that are technically sound and coherent.
 6. Implement a procedure as part of the achievement levels-setting process for selection of items to illustrate each level of achievement at each grade level and recommend those to the Board to be used for reporting results on the 2009 assessment. The Governing Board will specify the item selection criteria and procedure to be implemented.
 7. Collect and report information in support of the procedural validity of the recommended achievement levels. The Board wishes to have evidence of the procedural validity of the process and of the external validity of the outcomes. Evidence of external validity will be collected under a separate contract.

8. Produce a series of reports of the final recommended achievement levels, descriptions, and exemplar items and the methodological and technical procedures implemented to generate and support the validity of the recommended levels. These reports collectively shall provide detailed documentation of the science achievement levels-setting process, as well as recommendations for future achievement levels-setting efforts.
9. Provide reports on the progress of the contract to the Committee on Standards, Design and Methodology and to the Board when requested to do so by the COR.
10. Attend meetings and provide reports on the progress of contract work to select meetings of committees, panels, and advisory groups in support of the NAEP program when requested to do so by the CO under the technical guidance of the COR.

3.2 Procurement Tasks

To accomplish the purposes of this procurement, the contractor shall carry out the tasks outlined below and shall be expected to engage in technically sound, effective, and cost-efficient approaches to accomplish the setting of achievement levels on the 2009 science assessment.

3.2.1 Task 1: Develop Planning Document

Within seven business days following the award of the contract, the contractor's staff will meet with the CO, COR and other Governing Board staff in Washington DC to further define relationships and understand the work to be done. The meeting will also clarify the working relationships that will exist among the parties in accomplishing tasks identified in this scope of work. Within ten business days following this meeting, the contractor shall prepare and submit an electronic file version (in Microsoft Word 2003) of a written summary of the discussion, including all agreements reached for review and approval by the CO and COR. Comments on this document will be provided by Board staff within five business days following receipt.

Based on the schedule of activities discussed and the agreements reached at this initial meeting, the contractor shall prepare and provide to the CO, with a copy to the COR, a Planning Document which outlines by task, as well as chronologically for the contract period, each activity to be performed under this procurement in a level of detail that allows long-range planning for the Board and Board staff. Much of the information to be provided in this document will have been included in the proposal, but the details may be adjusted as a result of additional information provided in the initial meeting with Governing Board staff. This Planning Document shall also serve as a monitoring document for keeping the contract on schedule. The Planning Document shall be submitted within 30 calendar days of the award of the contract. Upon request by the COR, the contractor shall provide an updated version of the Planning Document to reflect changes in the schedule of activities. Ten printed copies of each version will be submitted with one copy in electronic format in Microsoft Word 2003.

3.2.2 Task 2: Design Achievement Levels-setting Procedures and Produce a Design Document

The contractor shall develop a comprehensive Design Document, which provides the rationale for the design of the achievement levels-setting activities for the 2009 science assessment. (See the table in Section 3.3 for a list of project deliverables.)

In developing the rationale for the design, the contractor shall follow the Board's policy on setting achievement levels. In addition to other factors, the design rationale must clearly reflect consideration of the following.

- Outcomes and issues related to the activities of the Board regarding the previous establishment of achievement levels;
- The unique design features of the 2009 science assessment (Please see the Science Framework and Test Specifications at www.nagb.org/frameworks for details);
- Relevant research and advances in the field of standard setting.

The designs shall be clear, creative and effective in the descriptions of panelists' tasks; the presentations of training, instructional, and informational materials; the displays of performance results and other feedback; the use of technology to analyze and report panelists' recommended levels; the collection of procedural validity evidence, and other components of the achievement levels activities.

The Design Document for the science ALS process shall include, but not be limited to, discussions of the following features of the process.

- **Achievement Levels-Setting Methodology** –The Governing Board has two previous experiences with the use of modifications of the Bookmark process for setting NAEP achievement levels. The Mapmark method using content domains was implemented for the 2005 NAEP grade 12 mathematics achievement levels-setting activity. The Mapmark procedure provides some improvements over the bookmark procedure and some advantages over the NAGB/ACT methodology previously implemented for NAEP achievement levels-setting activities. The Governing Board explored the applicability of this methodology for NAEP achievement levels setting in economics in a pilot study in 2006, and an additional modification called the Mapmark with Whole Booklets was selected for the operational achievement levels-setting process. (Please see Appendix B for locations of reports on www.nagb.org documenting the process implemented for economics.)

The offeror must design a procedure for setting achievement levels for science at grades 4, 8, and 12. This methodology must meet all requirements for NAEP achievement levels setting. The offeror must provide a rationale to explain the reasons why this procedure is appropriate.

A pilot study must be conducted to implement the exact procedure planned for the operational achievement levels-setting process. The offeror is encouraged to propose a clear, well organized, and cost efficient study design that maintains the independence of the outcomes of the panels both within and across grade levels. The size of panels for the pilot may be smaller than that proposed for the operational achievement levels-setting panels, but in all cases the panel size must be adequate to provide evidence of the technical soundness of the procedures implemented.

The offeror shall demonstrate an understanding of the procedures and describe the general plan for implementation of each element. The descriptions shall include an evaluation of the proposed methodology, significant citations in the literature, documented use by state(s) or other large-scale assessment programs, and other supporting information.

- **Achievement Levels Panels** – Procedures shall be outlined for obtaining broadly representative panelists to participate in the achievement levels activities. All panelists must be well qualified by content knowledge and by familiarity with the knowledge, skills, and abilities in science of students in the grade level assessed by NAEP for which the panelist is to serve as a member. Panels shall be composed of two-thirds (70%) educators and one-third (30%) non-educators and reflect an overall balance of gender, race/ethnicity, and geographic location. Classroom teachers currently engaged in science instruction at the grade level assessed by NAEP (4th, 8th, or 12th) should comprise 55% of the group. Other educators comprising 15% of the group should include individuals such as post-secondary science faculty teaching introductory courses, faculty in science teacher education programs, and state and local curriculum directors. Representatives of the general public should have some background or experience in science and with students in the general age and grade level of the assessment.

No honorarium has been paid to NAEP achievement levels-setting panelists in the past, although expenses in compliance with Federal Travel Regulations are paid and schools are reimbursed for the cost of substitute teachers. The offeror may propose payment of an honorarium or other incentives for panelists' participation, but justification must be provided to support the effectiveness of the incentive.

- **Briefing Materials** - The design document shall outline the types of background materials to be compiled and disseminated to the achievement levels participants for each procedure. These materials are intended to orient the panelists to the purpose of the activity; provide information on NAEP, The National Assessment Governing Board, and the science assessment; and provide the appropriate frame of reference.
- **Achievement Levels Setting** - The discussion of the actual levels-setting tasks shall outline the nature of the tasks and the procedures to be implemented. This will include the training of panelists; the contractor's key staff involved; the number, timing, and length of panel meetings; meeting agendas and evaluation forms; and the type and scheduling of information presented to panelists. Each levels-setting activity must include a segment in which panelists take a grade-level form of the science NAEP under conditions as similar as practicable to the NAEP administration procedures. Information on student performance relative to the achievement levels *must* be presented to panelists as feedback during the process, but the initial judgments must be based on the achievement levels definitions without data regarding actual student performance beyond that required to train panelists and implement a bookmark-type procedure with item maps. Previous experiences with NAEP achievement levels-setting activities must be *explicitly* taken into account in the design of each proposed process.
- **Data Processing** - The Design Document shall present plans for on-site data entry and analysis of data obtained from all phases of the ALS activities. Offerors are encouraged to take full advantage of the latest technology in the process of training and providing information to panelists and in the process of capturing judgments and other information from panelists. Plans shall also outline a strategy for examining and displaying results for presentation to the Committee on Standards, Design and Methodology and to the full Board.

- **Achievement Levels Descriptions and Sample Items** –The Governing Board shall conduct a process, outside the scope of this procurement, to finalize the achievement levels descriptions. The final descriptions to be used in the achievement levels process will be provided to the contractor by the Design Document due date. The contractor shall implement a procedure specified by the Governing Board for the selection of items to represent each achievement level, taking into account both the achievement levels definitions and the numerical cut scores associated with each level. The items are to be used to illustrate the knowledge, skills, and abilities in science attained by students performing at each level of achievement. The selection of exemplar items is a task of the achievement levels-setting panel and should be scheduled just after the final cut scores have been recommended. The exemplar items will be included in the evaluation of the outcomes of the ALS process. The offeror may propose ways to vet the set of exemplar items to recommend to the Governing Board.

- **Public Comment** – The contractor shall describe procedures for collecting public comment from a broad spectrum of groups and individuals concerned with the design of the science levels-setting procedures. In addition, the contractor is encouraged to design procedures to collect public comment regarding the outcomes of the levels-setting procedures. Comments shall be obtained from the education community—public and private—including state testing directors and science curriculum coordinators, and from key stakeholder groups in the business community. The contractor shall produce a detailed plan for announcing the public comment component to ensure a high rate of response. The information generated through collection of public comment may contribute to the evidence regarding the validity of the process.

The Design Document specified in this section shall be submitted to the National Assessment Governing Board through the CO for review, recommendations, and/or approval within 60 calendar days of the award of the contract.

3.2.3 Task 3: Conduct Achievement Levels-Setting Process

Task 3 is to implement the design developed in Task 2 in order to generate recommendations that will assist the Board in establishing achievement levels on the NAEP 2009 science assessment. The following timeline presents the projected dates for key activities associated with the 2009 science assessment and reporting of results. Please note that some dates may change; all dates are subject to the operational and reporting timelines established by NCES and the Board.

2009 NAEP Science Assessment Schedule

Activity	Projected Dates*
2008 Science Field Trial Data Available to Achievement Levels Contractor; Operational Item Pool Identified	October 2008
Operational 2009 Science Assessment Administration	February – March 2009
Scoring Item Data	April – June 2009
Scaling and Analysis of Data	June – August 2009
Data Available to Achievement Levels Contractor	August 2009
The Governing Board Sets Science Achievement Levels	March 5, 2010
Release The Nation’s Report Card for 2009 Science	June 2010

*These dates are approximate with each activity dependent on the previous ones. Exact dates will be finalized after NCES and the Board negotiate key dates with their contractors.

In planning and conducting the ALS work, the contractor shall coordinate all project tasks with the assessment schedule, allowing adequate time for review and decision-making by the COR, the Committee on Standards, Design and Methodology, and the Board. For planning purposes, the contractor should allow 30 calendar days for review and evaluation of the Design Document and results from any panel meetings approved for implementation in this contract. The contractor may schedule 10 working days for reviews of draft deliverables that only require input by the Governing Board COR.

- (A) The contractor shall identify specific groups and individuals to be involved in the achievement levels-setting process.
- (B) The contractor shall describe the roles of various participants in the achievement levels-setting process. The ALS process shall be designed to ensure the active involvement of relevant groups and individuals once they have been identified and asked to participate.
- (C) A set of public comment collection activities shall be developed that may include, but will not be limited to, meetings, canvassing of various groups and individuals, hearings, and written and oral communication that represent the contractor’s best effort to involve a broadly representative group in developing the achievement levels-setting activities and recommendations for the Governing Board.
- (D) The contractor shall be responsible for all plans and costs of logistical arrangements related to implementation of the proposed set of activities. Timelines for these activities shall be provided to the COR well in advance of their occurrence, and in accordance with the timelines specified in the approved Planning Document. Such arrangements include, but are not limited to the following activities.

1. Arranging for suitable meeting space
2. Making all travel arrangements and paying for or reimbursing expenses for meeting participants (exclusive of outside observers, evaluators, Board members, and Governing Board staff). Reimbursement of lodging and meals shall be in accordance with Federal Travel Regulations (see www.gsa.gov for current per diem amounts).
3. Reimbursing the cost of substitutes for teacher panelists
4. Developing agendas
5. Preparing briefing materials and disseminating the materials at appropriate times
6. Preparing and making available (when required by the COR) minutes or transcripts of all meetings and conferences within 7 working days following the relevant meeting or conference
7. Obtaining the NAEP test items and performance data used in the assessment from the NAEP Design, Analysis, and Reporting (DAR) contractor, via the Governing Board COR
8. Preparing all training materials and data collection forms
9. Conducting meetings for the achievement levels-setting activities
10. Entering, verifying, analyzing, and reporting all data resulting from the achievement levels activities
11. Preparing and presenting reports for COSDAM and the Board
12. Providing courtesy responses to all participants for the Board including preparation of individual letters of commendation and appreciation for the COR's signature
13. Preparing and distributing, with prior approval from the COR, press releases to the local newspapers of panelists to notify the local community of the contributions of the panelists.

(E) The contractor shall organize and structure the project to encourage effective participation of various groups and ensure appropriate timing of activities. Plans to accomplish this and pertinent timelines will be addressed in the offeror's proposal. It is critical that stakeholders be given the maximum opportunity to participate in the process.

(F) The contractor shall identify five (5) outside technical consultants to serve on the Technical Advisory Committee on Standard Setting (TACSS). The TACSS must include individuals with expertise in NAEP scaling and analysis procedures, and achievement levels-setting procedures. At least one TACSS member shall have been involved in at least one previous NAEP achievement levels-setting process for the Board. In addition, one member of the Design, Analysis, and Reporting contractor's staff who is directly involved with all operational procedures involving NAEP data scaling and analysis shall be the sixth member of TACSS. (ETS, Inc. is currently the Design, Analysis, and Reporting contractor. The COR will notify the achievement levels-setting contractor of the ETS representative to serve on the TACSS when the person has been named.) Names and affiliations of proposed TACSS members must be provided in the proposal for Governing Board approval.

While the contractor is required to engage a broad range of highly technical persons in advising on all technical issues that are raised during the period of this contract, all decisions, advice, and recommendations made to the Board shall be the sole responsibility of the contractor.

The contractor shall schedule an average of one in-person (one 2-day) meeting every 60 days during the planning and implementation phases of the contract. The contractor should fully inform TACSS and solicit effective advice on all technical aspects of the project.

- (G) The contractor is responsible for providing all numerical and other data necessary to calculate the final recommended achievement level cut scores and to identify items selected to illustrate each achievement level. The Design, Analysis, and Reporting contractor to NCES is ultimately responsible for ensuring that the data are mapped onto the NAEP scale.
- (H) One or more key staff of the contractor (to be specified by the COR) shall attend the Governing Board quarterly meetings and any special meetings of COSDAM or the Board that may be required to facilitate the goals of this contract. In addition, one or more key persons on the contractor's ALS staff may be asked to prepare and present reports on contract activities to other groups that advise the NAEP program.

For planning purposes, assume two key representatives of the contractor's project staff will attend all Governing Board meetings and make presentations regarding the activities of the project. Assume also that there will be at least two extra meetings of COSDAM when the contractor's key staff will fully brief the Committee on the results of the ALS process during a one-day meeting. Further assume that at least one representative of the contractor's project staff will attend a total of six additional meetings of various organizations and groups lasting an average of 2 days each in Washington, DC.

Task 3 must be completed in time for the resulting cut scores and illustrative items to be used for reporting the 2009 Science NAEP. The current plan is to report the results by June 2010. To meet this schedule, several targets must be met.

- First, the pilot study must be conducted and a report prepared for presentation to COSDAM by the May 2009 meeting of the Governing Board.
- The achievement levels panel must be convened and a report prepared for presentation to COSDAM in a focused briefing session in January 2010.
- An Interim Report of the recommendations resulting from the levels-setting activities and the collection of additional evidence shall be due at least 2 weeks prior to the March 2010 meeting of the Board to allow adequate time for Board input and decision-making regarding the achievement levels.
- The contractor must be prepared to recommend achievement levels to the Governing Board at the March 2010 quarterly Board meeting in order to have the report released in June 2010.

3.2.4 Task 4: Develop Final Reports

The achievement levels are the primary means of reporting student performance for *The Nation's Report Card*.™ Documentation of the process is a vital part of setting and reporting NAEP achievement levels. The quality and content of all reports prepared by the contractor is of utmost importance to the Governing Board. While other aspects of the final reports are deemed important, two requirements for the final reports must be fully met. The first requirement is to provide a clearly stated and well-organized documentation of the logistic, methodological, and technical aspects of the ALS process for the Board. The second requirement is to assure that these reports are of a quality and style that will yield information accessible to the broad

audience for NAEP achievement levels. The audience for the reports is the Governing Board, the education community, education policy makers, and the general public.

All final reports submitted by the contractor shall include an *executive summary*, the *full text and discussion*, and *appendices* containing all relevant tables and materials used in the ALS process such as written communications with participants, presentation slides, agendas, training materials, evaluation forms, as well as a listing of (a) members of the Governing Board; (b) members of the Committee on Standards, Design and Methodology; (c) the Governing Board staff assigned to the project; and (d) names of reviewers, as appropriate.

The executive summary and the full text and discussion components shall be written in a way that allows each to be presented as a standalone document suitable for public distribution separate from the appendices (for those who wish only a textual summary of the information). Ten printed copies and an electronic version (in Microsoft Word 2003 format) of all deliverables including both interim/draft and final versions, shall be provided to COR.

The contractor shall prepare two comprehensive final reports: a Process Report and a Technical Report.

The **Process Report** shall include, but not be limited to, the following:

- **Description of Achievement Levels-Setting Process** — The procedural aspects of the entire achievement levels-setting process, including the pilot study, shall be documented in a way consistent with providing evidence of procedural validity. The Process Report shall document the elements of the achievement levels-setting process and panelists' reactions to each. The extent to which this information provides evidence of procedural validity shall be evaluated and reported. This shall include data on panelists' ratings, changes in ratings, internal consistency of ratings, panelists' reactions to feedback data, and panelists' confidence in the process and satisfaction with the process and process outcomes. This report shall include descriptions of participants, training, meetings, tasks, materials, results, and final recommendations.
- **Recommended Achievement Levels** — Three recommended levels—basic, proficient, and advanced—shall be presented. All relevant data obtained during the ALS process must be aggregated for presentation to the Board, including appropriate summary statistics, estimates of reliability, and confidence intervals. Additionally, achievement level data for groups and subgroups included in NAEP reporting shall be included in briefings to the Board.
- **Achievement Levels Descriptions and Exemplars** — Clear and coherent descriptions of what students should know and be able to do in science at each achievement level for grades 4, 8, and 12 will be developed by the Board and provided for use throughout the achievement levels-setting process. Items selected according to the design to be specified by the Governing Board shall illustrate the achievement levels. The relationship and alignment to the achievement levels definitions and the scale score range shall be made explicit.
- **Public Comment** — Procedures and results related to obtaining public comment on the design of the process must be summarized for the Board in the final report.

- **Recommendations for Future Achievement Levels-setting Activities** — Based on the levels-setting activities conducted under this procurement and previous procurements, the contractor shall provide a thoughtful and reflective discussion of recommendations for future achievement levels-setting efforts.

The **Technical Report** shall include, but not be limited to, complete documentation of the ALS aspects delineated below. While the audience for the technical report may be somewhat more limited than that for the Process Report, the Technical Report must be clear, concise, and complete; and interested persons—including persons who are not trained in educational measurement—must be able to understand it.

- **Technical Advice** — All technical advice and decisions reached during this project, along with rationales and considerations for the decisions shall be documented and summarized. In addition, a complete set of minutes for the TACSS meetings shall be appended to the report.
- **Data Analysis Procedures** — All analysis techniques, decisions, formulas, and procedures must be documented clearly and completely. Original data forms, software created under the contract, and all raw data shall become the property of the Governing Board and shall not be released by the contractor to any other parties without written consent from the Governing Board. These data and software will be provided to the COR in electronic form, in addition to the written documentation.
- **Pilot Activities** — Technical aspects of the pilot activities conducted as part of the level-setting process shall be described and documented in this report including data collection activities, data analysis, and results.
- **Materials Analysis and Description** — All materials used throughout the process including briefing booklets, agendas, block pairings and block assignments to panelists, a tracking of item adjustments (e.g., items whose rubrics were changed from polytomous to dichotomous in the course of the scoring and scaling), and other relevant information shall be described and documented.

3.3 Schedule of Project Deliverables

For each deliverable, an electronic version and five draft copies shall be submitted to the COR for Board review before final submission. The Governing Board recommends that sufficient time be included in the schedule to consider feedback from the review of the draft in preparation of the final document. It is of paramount importance that the schedule for activities and deliverables be strictly adhered to and that the contractor allows sufficient time for Board review.

Two copies of final project deliverables identified in Section 3.2.4 including appendices, must be submitted on CD Rom to the Contracting Officer.

The following schedule provides a summary of the major project deliverables, and the due dates for each document. The Governing Board plans to award the contract no later than September 26, 2008.

2009 Science Levels-setting Activities: Project Deliverables

Document	Due Date
1. Planning Document	Within 30 days of contract award
2. Design Document	Within 60 days of contract award
3. Interim Process Report	March 2010
4. Interim Technical Report	May 2010
5. Final Technical Report	June 2010
6. Final Process Report	July 2010
7. Progress Reports	15 th of Each Month*

**The date can be based on the contractor's budget cycle, but it must be set for the same date each month.*

The contractor shall consult with the COR before making any major technical decision concerning a deliverable product. To facilitate this communication, the contractor shall have an electronic system for transferring information via facsimile machine and e-mail. This electronic system shall be fully accessible to staff seven days a week, 24 hours a day.

Monthly Reports

The contractor shall submit **monthly progress reports**, summarizing contractual activities and financial expenditures to the Board for the duration of the project. Monthly progress reports shall describe **contractual activities**, including major tasks and accomplishments, problems and suggested solutions, significant activities and events, any decisions that may be needed from the Board, and plans for the next reporting period. These progress reports shall be submitted by the fifteenth day of each month (or on a specified date agreed upon by the CO and the contractor for the preceding month). The CO, through the COR, shall respond within five calendar days to each monthly progress report with any technical direction or re-direction that may be warranted in order to carry out Board policy and direction.

In addition to monthly progress reports, the contractor shall provide monthly reports on **financial expenditures** through invoice submissions. Each invoice shall include the following information:

- (a.) A summary of the overall project costs broken down by the four task areas described in Section 3.2 Procurement Tasks. This will include task/contract funding; tasks costs by current reporting period; cumulative costs for each task; and a balance of funds for the overall contract.
- (b.) Following this summary, the contractor shall provide, by task, a breakdown of labor and other direct costs. Labor hours will be depicted by individual staff and/or consultant tallied by approved labor rates and labor costs for staff. Detail for other direct costs shall be summarized by line item description and cost. Budgeted versus actual costs for each line item will be provided. Sufficient detail for each cost, such as lodging and per diem, supplies, etc. shall be provided to substantiate billed costs.

A sample of the required invoice information is provided in Appendix C.

4. INSTRUCTIONS TO OFFERORS

4.1 Organization and Content of the Technical Proposal

Offerors are required to follow the proposal format and content suggestions detailed in this section and in Section L of the solicitation. Each offer shall consist of two separately packaged proposals: a technical proposal and a business proposal. All information necessary to judge the technical soundness and the management capabilities of the offeror will be contained in the technical proposal. The business proposal will contain all information related to the determination of the costs associated with each of the project's tasks.

The offeror shall also provide a Past Performance Report as part of the technical response to the Statement of Work. Section L of the solicitation provides information on the requirements for the Past Performance Report. This report shall consist of short abstracts of related work, for four previous projects completed during the past three years that identify clearly both the names of staff members who were participants and the name, current affiliation, and current telephone number of the sponsor's project officer. These project officers may be asked to report their experience with the bidder on relevant projects with regard to the size, problems (if any), cost overruns (if any), responsiveness, flexibility, and project quality. A form for filing the Past Performance Report is provided as Attachment B to the solicitation.

The technical proposal must not contain reference to specific costs, but resource information may be included so that the offeror's understanding of the scope of the work may be evaluated. The technical proposal shall be organized into the following sections: Table of Contents, Introduction; General Approach; Technical Work Plan; Management Plan; Related Experience of Proposed Staff; Related Organizational Experience (this includes, but is not limited to, the Past Performance Report prepared by the offeror); and References. Authors of each section shall be clearly identified in the proposal's Table of Contents. Specific requirements for each section are discussed in greater detail in the following paragraphs.

Before discussing the specific requirements for each section of the proposal, a special note is in order to clarify the general expectations for the offeror's technical proposal. The tasks and suggestions presented herein are intended to guide the prospective offerors. The successful proposal must be thorough and clear in the approach recommended. The proposal should fully describe the role and function of all groups involved. The schedule, however, is not flexible. The various components of the project must be completed by the stated dates because of the schedule of subsequent activities necessary to complete the analyses and to prepare and release the reports of the 2009 science NAEP assessment. The information presented is not intended to restrict an offeror's proposal or to stifle creativity. What is desired is the offeror's perspective on the approach and activities within the given schedule that can best accomplish the purpose of the achievement levels-setting process. It is on this assumption that the offeror should prepare a proposal as outlined below.

4.1.1 Table of Contents

The Table of Contents will provide an easy means to identify authors (if different authors are used to write various chapters of the proposal), and major points of discussion.

4.1.2 Introduction

The Introduction shall present the offeror's concept of the purposes, methodology, and products of the project and shall include a short summary of the offeror's qualifications and unique strengths related to project tasks as outlined in Section 3.

4.1.3 General Approach

The General Approach section shall present an overview of the general plan to accomplish the work successfully and provide the rationale for the proposed approach. In this section, the offeror is encouraged to demonstrate an understanding of key issues in standard setting and evaluate the extent to which these will be addressed by the proposed procedures. The merits of the proposed procedures and methodologies should be evaluated and compared to other standard setting methods used in NAEP.

4.1.4 Technical Work Plan

The Technical Work Plan shall provide a detailed discussion of how tasks outlined in Section 3 of this document are to be accomplished. The plan should expand upon each of the tasks and include a discussion of procedural and implementation issues related to completing each task. The comprehensive plan for carrying out each task shall identify staff members who will play a major role in its completion. Offerors are encouraged to propose efficient and effective approaches to accomplish the achievement levels-setting process.

Offerors are encouraged to propose the use of on-site or remote information processing capabilities to provide timely feedback to panelists (on their recommended achievement levels, for example), during the panel meetings. The successful offeror also shall provide samples of the types of feedback to be given to the panelists during the levels-setting process.

4.1.5 Management Plan

For a successful outcome, this project will require an effective management system that enables the contractor to complete tasks on schedule and within budget. The system shall include procedures for coordinating and controlling project personnel and tasks; ensuring adherence to schedules and deadlines; ensuring high quality products and outcomes; identifying potential problems early; maintaining close, effective communication with the COR; and accounting for and controlling project expenditures. The work of subcontractors (if any) is the responsibility of the prime contractor, and the management plan must specify how subcontractors will be managed.

The Management Plan shall detail the offeror's overall management plan, including proposed lines of authority, coordination and communication within the offeror's organization (or within a consortium, if applicable). The offeror shall describe in this section of the proposal the management control system that is used internally for planning, scheduling, budgeting, managing, and reporting the accomplishments of the contract.

Because of the nature of this project and extremely tight scheduling of activities, the project staffing patterns are especially important. The proposal shall state the name of each person assigned to work on the project, and identify the role and functions of each. One or more persons with expertise in the following areas may be proposed as project staff.

- Project management and direction
- Logistics planning and support
- Standard setting design, leadership and facilitation
- Surveying, sampling and statistical data analysis
- Psychometric analysis
- Operations support
- Written and verbal communications and presentations

The offeror is invited to propose a management and staffing plan that maximizes the efficiency of the process and assures adherence to the schedule of activities. The Governing Board stipulates, however, that a single person be identified as a project director to provide leadership and direction to the contractor's project staff and serve as the contractor's primary contact with the Governing Board management and technical staff and with the COSDAM and other Board members. The project director must participate in all aspects of the project, attend all meetings—both internal and external, and serve as the primary contact with and source of information for the contract. This person must have exceptionally good management, organizational, and communication skills.

The Governing Board recommends that a person with expertise in science assessment serve as an advisor to the contractor. The person(s) with expertise in science assessment may contribute through membership on TACSS, the project staff, or a review panel.

All personnel, including subcontractor staff and consultants, shall be identified and their positions in the contract's management structure detailed in an organization chart. This chart shall depict clearly the lines of authority and responsibility for all persons involved in the conduct of this project. Due to the number of presentations and written reports associated with this project, strong communications skills should be a priority for staffing the project. Writing samples (not to exceed 5 pages) for the project director and other key staff involved in report preparation and presentation should be submitted in an appendix to the technical proposal.

This section shall also include a Gantt chart depicting the timelines for all major tasks and subtasks, including deliverables, selection of panel members, panel meetings, the Governing Board meetings, and so forth. Included in this Gantt chart shall be the starting and completion dates for each task, as well as appropriate intermediate dates for precursor steps and draft deliverables, as appropriate. Staff responsible for each task shall also be included on the chart.

In all cases where personnel external to the offeror's organization are proposed, letters of agreement (with proposed staff) and letters of availability (endorsed by the primary institutions with which they are affiliated) shall be included.

4.1.6 Related Experience of Proposed Staff

The Related Experience of Proposed Staff section of the technical proposal shall consist of vitae for all proposed staff members. Primary areas of expertise should be clearly specified. To plan, conduct, and complete the work successfully, offerors must provide staff who together have technical expertise, knowledge, and experience in the following areas:

- Designing standard setting methodology
- Managing and coordinating standard setting efforts in large-scale assessments
- Developing large-scale assessments
- Communicating effectively both orally and in writing for a broad range of audiences
- Using negotiation skills and ensuring that the concerns of all stakeholder groups are addressed
- Coordinating efforts which involve a range of professional organizations, interests, and/or disciplines to achieve general agreement and understandings on outcomes among subject area and education constituencies

All staff shall have qualifications appropriate for the tasks they are to perform. Vitae of proposed staff shall document that the person has the knowledge, skills, and abilities associated with the tasks to be performed. The documentation shall include directly related experience, educational background, publications and roles played in related projects.

For each key staff member, identify two or three projects for which the person has played a major role and provide contact information (name, current affiliation, email address, and telephone number) for the sponsor's project officer or other staff member who is familiar with the key staff member's performance on each project.

Offerors shall submit performance indicators for all key project staff, particularly those who will lead and train panelists in the levels-setting activity (e.g., recording of presentations to large groups, evaluations, and so forth).

The proposed project director must have demonstrated project management skills that include successful project planning and execution, with effective contingency planning. Successful oversight for project cost controls is essential to the contract, and the project director must have demonstrated success. In addition, the project director and other key staff must have demonstrated ability to meet deadlines, to produce high quality products within the budget, and to keep the COR informed of and involved in major decisions or events that are likely to affect the project's performance or products. Any changes or substitutions of the key personnel will require written advance approval by the Contracting Officer.

4.1.7 Related Organizational Experience

The Related Organizational Experience section shall describe the offeror's past experience in conducting projects of a similar nature and magnitude. The organization or consortium of organizations undertaking this project must have proven capabilities in the following areas:

- Designing, implementing, and managing a large-scale standard setting process
- Working closely with national, state, and local education leaders, organizations, practitioners state and local school administrators, teachers, assessment specialists, researchers, state and local curriculum specialists, and organizations involved in K-12 education.

In addition, the organization undertaking the work must have performed well on prior contracts by demonstrating:

- Low staff turnover
- High quality control standards
- Adherence to budget limitations
- Responsiveness to the procurement project director
- Timeliness and acceptability of project deliverables

Short abstracts of related work shall be included which clearly identify both the names of staff members who were participants and the name, current affiliation, and current telephone number of the sponsor's project officer. These project officers may be asked to report their experience on relevant projects including factors such as the size of the project, any problems, any time or cost overruns, the responsiveness of staff, flexibility, and project quality.

4.2 Business Proposal

The Business Proposal shall be prepared in accordance with the solicitation requirements in Section L and shall include detailed labor cost information reported *by task* for all staff, consultants, and subcontractors assigned to work on the contract. In addition, the cost of labor per hour shall be provided by task for each staff assigned to work on the contract. Other costs such as airfares, lodging, office supplies, and reproduction costs should be separately identified by task. The offerors shall budget lodging and meal costs at government per diem rates posted on www.gsa.gov. The successful bidder must present an overall budget for the entire project with separate budget breakdowns by task.

Expenditures for TACSS meetings and attendance at quarterly Governing Board meetings should be allocated to the task scheduled to be the primary activity of the contract during that time period. For the early tasks with specific timeline requirements, TACSS meetings will be assigned to tasks according to the schedule of activities.

Offerors shall include sufficient costs to cover the attendance of key project staff at Governing Board quarterly meetings (see Appendix D for meeting schedule and locations) during the course of the contract, two additional one-day meetings of COSDAM anticipated for review of the outcomes of the panel meetings, and six additional meetings (to be determined by the COR) in Washington DC.

4.3 Use of Subcontractors

Proposals may include plans to subcontract parts of the work, provided evidence is presented that any proposed subcontractor has agreed to participate and is fully capable of performing the assigned tasks and that the offeror will have effective control of the subcontractor's work on the project. Offerors are encouraged to make use of subcontractors for specialized tasks where subcontractor expertise would strengthen the offeror's proposal. Contracts that include subcontractors shall be executed in accordance with the requirements of the prime contract, which is proposed as a cost plus fixed fee contract. The subcontractors shall submit detailed cost proposals adhering to similar requirements described in the Request for Proposals. Such proposals may, if the contractor desires, submit the cost proposals separately to the Governing Board. Invoice submission by the subcontractors shall also provide similar cost detail as described in Section 2.6 of this Work Statement.

Offerors are encouraged to make use of subcontractors for specialized tasks where subcontractor expertise would strengthen the offeror's proposal. Offerors may wish to subcontract with a professional logistical support service agency for securing hotels and contracting for food and other hotel services, supplying meeting agendas and briefing books, booking airline reservations, arranging ground transportation, reimbursing panelists, and so forth. Offerors are especially encouraged to subcontract logistic support services if such staff are not *directly* employed by the offeror. If a travel agency is proposed by the offeror, the proposal shall contain clear *evidence* that the agency can and will supply the services required by this contract.

4.4 Period of Performance

The period of performance will be 24 months. The contractor will be held accountable for meeting the due dates for all project milestones and all project deliverables indicated in the Scope of Work section of this solicitation.

APPENDICES

Appendix A - National Assessment Governing Board Achievement Levels Policy and Implementation Guidelines

Appendix B - Website Locations of Resource Materials for NAEP Achievement Levels-Setting and NAEP Science

Appendix C - Sample Invoice

Appendix D - Schedule of Board Meetings for 2008–2010

Adopted: March 4, 1995



APPENDIX A: Achievement Levels Policy and Implementation Guidelines

Developing Student Performance Levels for the National Assessment of Educational Progress

Policy Statement

Foreword

A policy on setting achievement levels on the National Assessment of Educational Progress (NAEP) was first adopted in 1990 and amended several times thereafter. The present policy, adopted in 1995, contained introductory and explanatory text, principles, and guidelines. Since 1995, there have been several changes to the NAEP authorizing legislation (currently, the NAEP Authorization Act: P.L. 110-279). In addition, related legislation has been enacted, including the No Child Left Act of 2001. Consequently, introductory and other explanatory text in the original version of this policy, no longer germane, has been deleted or revised to conform to current legislation. The Principles and Guidelines remain in their original form except for Principle 4, from which the reference to the now decommissioned Advisory Council on Education Statistics has been deleted. (Foreword added August 2007.)

Principles for Setting Achievement Levels

Principle 1

The level setting process shall produce for each content area, three threshold points at each grade level assessed, demarcating entry into three categories: *Basic, Proficient, and Advanced.*

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.

Principle 2

Developing achievement levels shall be a widely inclusive activity of the Board, utilizing a national consensus approach, and providing for the active participation of teachers, other educators (including curriculum specialists and school administrators at the local and state levels), and non-educators including parents, members of the general public, and specialists in the particular content area.

The development of achievement levels shall be conducted in two phases. In phase 1, the assessment framework development process shall yield preliminary descriptions of the achievement levels (*Basic, Proficient, and Advanced*), which shall subsequently be used in phase 2 to develop the numerical standards (cut scores) and to identify appropriate examples of assessment exercises that typify performance at each level. The levels will be updated as appropriate, typically when the assessment frameworks are updated.

Principle 3

The Governing Board shall incorporate the student performance levels into all significant elements of NAEP, including the subject area framework development process, exercise development and selection, and the methodology of the assessment. The achievement levels shall be used to report the results of the NAEP assessments so long as such levels are reasonable, valid and informative to the public.

Principle 4

In carrying out its statutory mandate, the Governing Board will *exercise its policy judgment in setting the levels*. The Board shall continually seek better means of setting achievement levels. In so doing, the Board may seek technical advice as appropriate from a variety of sources, including external evaluations provided by the Secretary, the Commissioner, and other experts. Proposed achievement levels shall be reviewed by a broad constituency, including consumers of NAEP data, such as policymakers, professional groups, the states and territories. In carrying out its responsibilities, the Board will ordinarily engage the services of a contractor who will prepare recommendations for the Board's consideration on the levels, the descriptions, and the exemplar exercises.

Guidelines for Setting Achievement Levels

Each guideline presented below is accompanied by a rationale and a summary of the implementation practices and procedures to be followed in carrying out the principle. It should be understood that the full implementation of this policy will require the contractor, through Governing Board staff, to provide assurances to the Board that all aspects of the practices and procedures for which they are responsible have been completed successfully. These assurances will be in writing, and may require supporting documentation prepared by the contractor and/or Governing Board staff.

Summary of Guidelines

Guideline 1

The level setting process shall produce for each content area, three threshold points at each grade level assessed, demarcating entry into three categories: *Basic, Proficient, and Advanced*.

Guideline 2

The level setting process shall be a widely inclusive activity of the Board, carried out by a broadly representative body of teachers, other educators (including curriculum specialists and local and state administrators), and non-educators including parents, concerned members of the general public, and specialists in the particular content area; this process and resulting products shall be reviewed by a broad constituency.

Guideline 3

The level-setting process shall result in achievement level cut scores for each grade and level, expanded descriptions of the content expected at each level based on the preliminary descriptions provided through the national consensus process, and exemplar exercises that are representative of the performance of examinees at each of the levels and of the cognitive expectations for each level described.

Guideline 4

In carrying out its statutory mandate, the Board will *exercise its policy judgment in setting the levels*. However, in so doing, they will seek technical advice from a variety of sources, but especially from the contractor who will prepare the recommendations on the levels, the descriptions, and the exemplar exercises, as well as from consumers of NAEP data, including policymakers, professional groups, the states, and territories.

Guideline 5

The achievement levels shall be the initial and primary means of reporting the results of the National Assessment of Educational Progress at both the national and state levels.

Guideline 6

The level-setting process shall be managed in a technically sound, efficient, cost-effective manner, and shall be completed in a timely fashion.

Guideline 1

The level setting process shall produce for each content area, three threshold points at each grade level assessed, demarcating entry into three categories: *Basic*, *Proficient*, and *Advanced*.

Rationale

The Board is committed to describing the full range of performance on the NAEP scale, for students whose performance is in the mid-range, as well as for those whose performance is below and above the middle. It is highly desirable to endorse realistic expectations for all students to achieve no matter what their present performance might be. Three benchmarks on the NAEP scale suggest realistic expectations for students in all regions of the performance distribution. Likewise, the Board is committed to preserving trend results in NAEP. Three achievement levels accommodate growth (and possible declines) in all ranges of the performance distribution.

Practices and Procedures

Policy Definitions

The following policy definitions will be applied to all grades, 4, 8, and 12, and all content areas in which the levels are set. It is the Board's view that the level of performance referred to in the policy definitions is what students *should be able to know and do*, and not simply the current academic achievement of students or that which today's U.S. schools expect.

- 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.

From Policy Definitions to Content Descriptions

In the course of applying the policy definitions to the level-setting process, it will be necessary to articulate them in terms of the specific content and sequence (now called descriptions) appropriate for the grades in which the levels are being set. This will be completed on a preliminary basis through the process which develops the assessment frameworks. These preliminary descriptions will be used to initially guide the work of deriving the advice that will assist the Board in setting the levels. Throughout the process of obtaining such advice, however,

these descriptions may be refined, expanded, and edited to more clearly reflect the specific advice on the levels.

Training of Judges

In training the judges for the level-setting activity, it is necessary that all arrive at a common conceptualization of *Basic, Proficient, and Advanced* based on the policy definitions of the Board. Such conceptualizations must be within the scope of the assessment framework under consideration and capable of being applied at the individual item level (Reid, 1991.)

Judges must also be trained in the specific model that will be used to generate the rating data. At the very least, they need to understand the purposes for setting the levels, the significance of such an activity, the NAEP assessment framework for the subject area under discussion, elements that make particular exercises more or less difficult, and the rating task itself.

Judges shall be trained by individuals who are both knowledgeable in the subject matter area and are experienced, capable trainers in a large-group setting. Presentations shall be prepared, rehearsed, and piloted before implementation.

Judges shall be provided comprehensive, user-friendly training materials, adequate time to complete the task, and the appropriate atmosphere in which to work, one that is quiet, pleasant, and conducive to reaching the goals of the level-setting activity. It is also required that judges take the assessment under the same NAEP-like conditions as students, that is, using the NAEP student booklets, having all manipulatives and ancillary materials, and timed.

Guideline 2

The level setting process shall be a widely *inclusive* activity of the Board, carried out by a broadly representative body of teachers, other educators (including curriculum specialists and local and state administrators), and non-educators including parents, concerned members of the general public, employers, scholars, and specialists in the particular content area. This process and resulting products shall be reviewed by a broad constituency.

Rationale

The spirit of the legislative mandate of the Board is one of moving toward a national consensus on policy issues affecting NAEP. The Board has historically involved broad audiences in its deliberations. The achievement levels are no different. Further, the Board views the level-setting activity as an extension of the widely inclusive effort to derive the assessment frameworks and scope and sequence of each assessment. Finally, the magnitude of the decisions regarding *what students should know and be able to do* is simply too important a decision to seek involvement from professionals alone; it must have the benefit of the collective wisdom of a broadly representative body, educators and non-educators alike.

Practices and Procedures

Sample of Judges

The panel of judges will be composed of both educators and non-educators. About two-thirds of the panel will represent teachers and other educators; one-third will represent the public, non-educator sector, for example, scholars, employers, parents, and professionals in occupations related to the content area. They will be drawn from a national sampling frame and will be broadly representative of various geographic regions (Northeast, Southeast, Central, West, and the territories) types of communities (urban, suburban, rural), ethnicities, and genders.

Individual panel members shall have expertise in the specific content area in which the levels are being developed, expertise in the education of students at the grades under consideration, and a general knowledge of assessment, curriculum, and student performance. The composition of the panels should be such that they meet the requirements of the *Standards (1985)*.

The size of the panels should be responsive to what the research demonstrates regarding numbers of judges involved (see Jaeger, 1991). While it may not be practical or beyond the resources available, every effort should be made to empanel a sufficient number of judges to reduce the standard error of the cut score. While there is no absolute criterion on the magnitude of the standard error of the cut score, a useful rule of thumb is that it should not exceed the *combined* error associated with the standard error of measurement on the assessment and the error due to sampling from the population of examinees.

Review Procedures

Throughout the process and particularly at critical junctures, groups that have a legitimate interest in the process will be involved. During the planning process interested groups and individuals will be encouraged to participate and share their experiences in the area of setting standards. These groups might include professional societies, *ad hoc* advisory groups, standing advisory committees to the Governing Board or its contractor(s) and NCES and its contractor(s) and grantees. Documents (such as the Design Document and Interim Reports) will be disseminated in sufficient time to allow for a thoughtful response from those who wish to provide one.

Proposed levels will be widely distributed to major professional organizations, state and local assessment and curriculum personnel, business leaders, government officials, the Planning and Steering Committees of the framework development process, the Exercise Development panels, and other groups who may request them.

When it is deemed useful by the Board, public hearings and forums will be conducted in Washington, D.C. and other parts of the country to encourage review and input on a broad regional and geographic basis.

Guideline 3

The resulting products of the level-setting process shall be (1) achievement level scores marking the threshold score for each grade and level, (2) expanded descriptions of the content expected at each level based on the preliminary descriptions provided through the national consensus process, and (3) exemplar exercises that are representative of the performance of examinees at each of the levels and of the cognitive expectations for each level described. These three products form the basis for reporting the results of all future NAEP assessments.

Rationale

The NAEP scale, while useful for aggregating large amounts of information about student performance in a single number, requires contextual information about the specific content and the sequencing of that content across particular grades, in order to be truly beneficial to users of NAEP data. In order to make the NAEP data more useful, descriptions of each level which articulate content expectations and exemplar exercises taken from the public release pool of the most current NAEP assessment must accompany the benchmarks or cut scores for each level. The descriptions and exemplars are intended to be illustrative of the kind of content that is represented in the levels, as well as an aid in the interpretation of the NAEP data.

Practices and Procedures

Methodology

The methodology to be used in generating the levels will depend upon the specific assessment formats for the content area in which the levels are being set. Historically, in the case of multiple choice exercises and short constructed response formats, a modified Angoff (1971) procedure has been employed. In the case of extended constructed response formats, a paper-selection procedure has been employed. Neither of these is without its disadvantages. As the assessment formats of future assessments become more complex and employ more performance-type exercises, it is quite likely that alternate procedures will be needed. The Board will decide these on a case-by-case basis, looking for advice from those who have had experience in dealing with these alternative assessment formats. In any case, the design for carrying out the process must be carefully crafted, must be appropriate to the content area and philosophy of the assessment framework, and must have a solid research base.

The procedures will generally be piloted prior to full implementation. The purpose of the pilot would be to test out the materials used with the judges, the training procedures, the feedback information given to the judges during the process, and the software used to complete the initial analyses. Procedures would be revised based on the pilot experience and evaluation evidence.

Whatever methodology is used, all aspects of the procedures will be documented for the purposes of providing evidence of procedural validity for the levels being recommended. This evidence will be made available to the Board at the time of deliberations about the levels being set.

Quality Control Procedures

While there are numerous points in a complex process for mistakes to occur, there are at least three important junctures where quality control measures need to be in place. First, is the point of data entry. Ideally, judges' ratings should be scanned to reduce manual errors of entry. However, if the ratings are entered manually, then they shall be entered and 100% verified using a double-entry, cross-checking procedure. Second, software programs designed to complete initial analyses on the rating data must be run with simulated data to de-bug, and provide assurances of quality control. The programs should detect logical errors and other kinds of problems that could result in incorrect results being generated. Finally, the production of cut scores on the NAEP scale is the final responsibility of the NAEP operations contractor. Only final cut scores, mapped onto the properly weighted and equated scale, received in writing from the operations contractor, will be officially communicated to the Board, or others who have a legitimate need to know. *Once the accuracy of the data has been ensured by the level-setting and operations contractors, the Board shall make a policy determination and set the final achievement levels, informed by the technical process of the level-setting activity.*

Descriptions of the Levels

The preliminary descriptions developed through the framework development process will be the starting point for developing recommendations for the levels under consideration. The preliminary descriptions are *working descriptions* for the panels while doing the ratings. These may be expanded and revised accordingly as these panels conduct the ratings, examine empirical performance data, and work to develop their final recommendations on the levels. The recommended descriptions will be articulated in terms of what students *should know and should be able to do*. They shall be coherent within grade, and consistent across grades, and will reference performance within the three regions created by the cut scores. No descriptions will be done for content below the *Basic* level.

Exemplar Exercises

The exemplars chosen from the released pool of exercises for the current NAEP assessment will reflect as much as possible performance both in the *Basic, Proficient, and Advanced* regions of the scale, as well as at the threshold scores. Exemplars will be selected to meet the $rp = .50$ criterion, and will demonstrate the range of performance possible within the regions. They will likewise reflect the content found in the final descriptions and the range of item formats on the assessment. Evidence will be provided for the degree of congruence between the content of the exemplars and that of the descriptions. There will be at least three exemplars per level per grade identified.

Guideline 4

In carrying out its statutory mandate, the Board will *exercise its policy judgment in setting the levels*. However, in so doing, they will seek technical advice from a variety of sources, but especially from the contractor, who will prepare the recommendations on the levels, the descriptions, and the exemplar exercises, as well as from consumers of NAEP data, including policymakers, professional groups, the states and territories.

Rationale

Setting achievement levels is both an *art* and a *science*. As an *art*, it requires judgment. It is the Board's best policy judgment what the levels should be. However, as a *science*, it requires solid technical advice based on a sound technical process. The Board is committed to seeking such technical advice from a variety of sources.

Practices and Procedures

Technical Advice throughout the Process

The Board seeks to involve persons who have had experience in standard-setting at the state level, and from those who are users of the NAEP results. Regular presentations will be given to standing committees who advise on NAEP matters such as the Education and Information Advisory Committee (EIAC) of the CCSSO, and the NAEP NETWORK. Their counsel will be sought on matters of substance as the work of the Board progresses. The EIAC and other similar constituencies may also be invited to send a representative to all standing technical advisory committees of the Board's contractor(s) which deal with the level-setting process.

The Board will also seek advice from the technical community throughout the level-setting process. Efforts will be made to ensure that presentations are made regularly to such groups as the American Educational Research Association (AERA), the National Council for Measurement in Education (NCME), and the professional groups in the content areas such as the International Reading Association (IRA), the National Science Teachers Association (NSTA), and other similar organizations. The Board will seek to engage technical groups available to them, including the Technical Review Panel, the National Academy of Education, their own contractor(s), and NCES and its contractor(s), in constructive research studies focused on providing information on the technical aspects of NAEP related to level-setting (e.g., scaling, weighting, mapping ratings to the scale, etc.)

Validity and Reliability Evidence

The Board will examine and consider all evidence of reliability and validity available. These data would include, but need not be limited to, procedural evidence such as the selection and training of judges and the materials and methods used in the process, reliability evidence such as intra-judge and inter-judge consistency data, and finally, internal and external validity data. Such data will help to inform *the Board's policy decision as they set the levels*.

Procedural evidence, while informative, is not necessarily sufficient evidence for demonstrating the validity of the levels. Therefore, the conduct of the achievement level-setting process shall be implemented so that a series of both internal and external validation studies shall be conducted simultaneously. To the extent possible, in order to realize maximum efficiencies in the use of resources, validation studies shall be included in the design of the level-setting data collection activities. Such studies may include, but shall not be limited to, convergent and divergent validation efforts, for example, conducting alternate standard-setting methods or conducting cross-validation level-setting activities, as well as exploring alternate methods for refining and expanding the preliminary achievement levels definitions, and empirically examining various technical decision rules used throughout the process.

As part of the validation task, additional evidence as to the suitability and appropriateness of identifying the subject area content of the recommended achievement levels ranges and cut-scores will be gathered. This evidence may include, but need not be limited to, data resulting from behaviorally anchoring the ranges and/or cut-scores, or data resulting from some other alternative procedures that employ a more global approach other than the item content of the particular assessment. The results of these studies will provide a clear indication of what students know and can do at the levels.

The results from these 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. Kane (1993) suggests that an “interpretive argument would specify the network of inferences leading from the score to the conclusions drawn about examinees and the decisions made about examinees, as well as the assumptions that support these inferences.” An interpretative argument which articulates the rationale for interpreting the levels shall accompany the presentation of proposed levels to the Board.

Again, to maximize the efficient use of resources and to minimize duplication of effort, it is highly desirable for contractors to coordinate the design of such studies with other agencies responsible for evaluating the level-setting activities.

Guideline 5

The achievement levels shall be the initial and primary means of reporting the results of the National Assessment of Educational Progress at both the national and state levels.

Rationale

In an effort to improve the form and use of NAEP the Board seeks to make the results of NAEP more accessible and understandable to the general public and to policy makers. The Board also supports the movement from norms-based assessments to standards-based assessments. Reporting the results of NAEP using the achievement levels accomplishes these ends to a greater degree than heretofore possible.

Practices and Procedures

Reporting What Students Know and Can Do

The purpose of most NAEP reports, but particularly those published under the auspices of the National Center for Education Statistics, is to report to the American public and others on the performance of students—that is, to report on *what students know and can do*. The purpose of the achievement levels is to identify for the American public what students *should know and should be able to do*, and to report the actual performance of students in relation to the achievement levels. Therefore, NAEP reports incorporate elements of both of these aspects of performance.

Clarity of interpretation of the NAEP data can be achieved by ensuring that the descriptions of performance for the levels and the exemplar exercises reflect what the empirical data show for a given assessment. This may be achieved by the modified procedures of *scale*

*anchoring*¹ or by new procedures developed specifically for the purposes of providing elements of the content of the frameworks in the reporting mechanisms.

Reporting Student Performance

In describing student performance using the levels, terms such as *students performing at the Basic level* or *students performing at the Proficient level* are preferred over *Basic students* or *Proficient students*. The former implies that students have mastery of particular content represented by the levels, while the latter implies an inherent characteristic of individual students.

In reporting the results of NAEP, the application of the levels of *Basic*, *Proficient*, and *Advanced* applies to the three regions of the NAEP scale generated when the appropriate cut scores are mapped to the scale. However, three cut scores yield, in fact, four regions. The region referenced by content which falls below the *Basic* cut score will be identified by descriptors that are not value-laden.

Interpreting Student Performance

When interpreting student performance using the levels, one must diligently avoid over interpretations. For example, each of the NAEP subject areas are scaled independently of each other, even though each scale uses the same metric, i.e., scores ranging from 0 to 500. Because the metrics are identical, it does not follow that comparisons can be made across subjects. For example, a *Proficient* cut score of 235 in reading should not be interpreted to have the same meaning as a *Proficient* cut score of 235 in U.S. history. Neither should unwarranted comparisons be made in the same subject area from one assessment year to the next, unless the data for the two years have been equated and we have reason to believe that the scale itself has not changed from time 1 to time 2.

Guideline 6

The level-setting process shall be managed in a technically sound, efficient, cost-effective manner, and shall be completed in a timely fashion.

Rationale

Since a contractor(s) is conducting technical advisory and assistance work for the Board, it is critical that such work be performed to meet high quality standards, including efficiency, cost-effectiveness, timeliness, and adherence to sound measurement practices. *However, in the final analysis, it is the Governing Board that makes the policy decision regarding the levels, not the contractor.*

Practices and Procedures

The contractor(s) shall prepare a fully detailed Planning Document at the onset of the level-setting work. This document will guide the progress of the work, serve as a monitor, and be the basis for staff and Board supervision. The Planning Document will outline milestone events in the process, provide a chronology of tasks and subtasks, as well as a monthly chronology of

all activities across all tasks, and detail all draft and final documents that will be produced, the audience for such reports, and the number of copies to be provided by the contractor.

Procedures adopted by a contractor(s) to carry out the level-setting process must encourage and support national involvement by the relevant and required publics. Such meetings will also be conducted in a physical environment which is conducive to work and planning. To the extent possible, current technology shall be used in all areas of the level-setting process to increase efficiency and to reduce error.

The contractor(s) shall work closely and in a professional manner with the NAEP operations contractor in striving to fulfill the requirements of the level-setting process by (1) making all requests for information and data in a timely manner, (2) providing all requested information and data in a timely manner, (3) adhering to all predetermined deadlines so as not to impede the work of the operations contractor, and (4) advising the operations contractor of all unusual findings in the data so that a concerted effort can be mounted to resolve the problem or issue at hand.

The contractor(s) shall develop the initial level-setting design adhering to sound measurement principles and ensure that the various components of the design (e.g., selection of judges) are congruent with current standard-setting research. In the implementation of such designs, they shall employ state-of-the-art training strategies and measurement practices.

The contractor(s) shall produce documents in a timely manner and make oral presentations upon request. Presentations may include, but need not be limited to, the Board's quarterly meetings, relevant Board committees, and professional and lay groups.

References

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (1985). *Standards for Educational and Psychological Testing*. Washington, DC: APA.
- Angoff, W.H. (1971). Scales, norms, and equivalent scores. In R.L. Thorndike (Ed.), *Educational Measurement (2nd ed., pp. 508-600)*. Washington, DC: American Council on Education.
- Jaeger, R.M. (1991). Selection of judges for standard-setting. *Educational Measurement: Issues and Practice, 10*, 3-6, 10, 14.
- Kane, M. (1993). The validity of performance standards. Unpublished manuscript.
- National Academy of Education (1992). *Assessing student achievement in the states. The first report of the National Academy of Education panel on evaluation of the NAEP trial state assessment: 1990 trial state assessment*. Stanford, CA: Author.
- National Assessment of Educational Progress Authorization Act, (P.L. 110-279).
- Reid, J.B. (1991). Training judges to generate standard-setting data. *Educational Measurement: Issues and Practice, 10*, 11-14.

Endnotes

1. The traditional scale anchoring procedures anchored at the 200, 250, 300 350 points of the scale (± 12.5 points), using a $p = .65$, and a discrimination of $.30$ with the next lower level. The modified anchoring procedures (tried in reading for 1992) anchored at the achievement levels cut scores (± 12.5), using a $p = .65$, and no discrimination criterion.

APPENDIX B: Website Locations of Resource Materials

**Developing Achievement Levels
On The 2009 National Assessment Of Educational Progress
For Science at Grades 4, 8, and 12**

**PUBLIC LAW 107-279, TITLE III: THE NATIONAL ASSESSMENT OF EDUCATIONAL
PROGRESS, SECTION 302: THE NATIONAL ASSESSMENT GOVERNING BOARD**

Under The Law in NAEP tab at www.nagb.org

SCIENCE 2009 FRAMEWORKS AND TEST SPECIFICATIONS

Under **Frameworks** at www.nagb.org

THE NATION'S REPORT CARD FOR 2005 SCIENCE

<http://nces.ed.gov/nationsreportcard/science>

**DEVELOPING ACHIEVEMENT LEVELS ON THE 2006 NATIONAL ASSESSMENT OF
EDUCATIONAL PROGRESS IN GRADE TWELVE ECONOMICS: TECHNICAL REPORT**

Click on document for link to www.nagb.org

**DEVELOPING ACHIEVEMENT LEVELS ON THE 2006 NATIONAL ASSESSMENT OF
EDUCATIONAL PROGRESS IN GRADE TWELVE ECONOMICS: PROCESS REPORT**

Click on document for link to www.nagb.org

**DEVELOPING ACHIEVEMENT LEVELS ON THE 2005 NATIONAL ASSESSMENT OF
EDUCATIONAL PROGRESS IN GRADE TWELVE MATHEMATICS: TECHNICAL REPORT**

Click on document for link to www.nagb.org

**DEVELOPING ACHIEVEMENT LEVELS ON THE 2005 NATIONAL ASSESSMENT OF
EDUCATIONAL PROGRESS IN GRADE TWELVE MATHEMATICS: PROCESS REPORT**

Click on document for link to www.nagb.org

The Technical and Process Reports referenced above are currently not available on the NAGB web site. A notice will be posted on the Fedbizopps web site when they are available:

<https://www.fbo.gov/index?s=opportunity&mode=list&tab=list>

APPENDIX C: Sample Invoice

SUMMARY PAGE

Two original copies of invoices need to be submitted to:

Ms. Remona Flowers
National Assessment Governing Board
800 North Capitol Street NW, Suite 825
Washington, DC 20002

Insert Contractor Contact information for invoice questions:

Contract # ED08XXXXXXXXXX

Period of Performance:

Billing Period:

Direct Labor:

Other Direct Costs:

Description	Contract Budget	Current Expenditures	Cumulative Expenditures	Contract Balance	% of Budget Expended
	<i>(Illustrative numbers)</i>				
Task 1 Develop Planning Document	\$100,000	\$5,000	\$20,000	\$80,000	20.00%
Task 2 Design ALS Process and Produce Design Document	\$100,000	\$5,000	\$20,000	\$80,000	20.00%
Task 3 Conduct ALS Process	\$100,000	\$5,000	\$20,000	\$80,000	20.00%
Task 4 a Validation Study - Booklet Classification Study	\$80,000	\$5,000	\$20,000	\$60,000	25.00%
Task 4 b Validation Study - Item Classification Study	\$90,000	\$5,000	\$20,000	\$70,000	22.22%
Task 5 Develop Final Reports	\$20,000	\$5,000	\$20,000	\$0	100.00%
Total Contract	\$490,000	\$30,000	\$120,000	\$370,000	24.49%

EXAMPLE OF REPORT BY TASK

Contract # ED08XXXXXXXXXX

Period of
Performance:
Contract #:

Billing Period:
Direct Labor:
Other Direct Costs:

Task Order 1: Develop Planning Document

Personnel/Direct Labor	Budget Hours	Labor Rate	Current	Cumulative	Total	Contract	Contract
			Labor Hours	Labor Hours	Task Costs	Task Budget	Task Balance
Person 1	40	\$120.00	8	16	\$1,920.00	\$4,800.00	\$2,880.00
Person 2	40	\$60.00	8	16	\$960.00	\$2,400.00	\$1,440.00
Person 3	8	\$60.00	8	8	\$480.00	\$480.00	\$0.00
Person 4	8	\$60.00	8	8	\$480.00	\$480.00	\$0.00
Person 5	8	\$60.00	8	8	\$480.00	\$480.00	\$0.00
Person 6	8	\$60.00	8	8	\$480.00	\$480.00	\$0.00
Person 7	8	\$60.00	8	8	\$480.00	\$480.00	\$0.00
Person 8	8	\$60.00	8	8	\$480.00	\$480.00	\$0.00

Total Costs	128	540	64	80	\$5,760.00	\$10,080.00	\$4,320.00
--------------------	------------	------------	-----------	-----------	-------------------	--------------------	-------------------

Other Direct Costs (ODCs)	Quantity	Unit Cost	Total
Supplies		\$100.00	\$0.00
Postage		\$100.00	\$0.00
Travel			\$0.00
Airfare	4	\$1,000.00	\$4,000.00
Per diem	4	\$64.00	\$256.00
Transportation	4	\$75.00	\$300.00
Consultants	2	\$1,500.00	\$3,000.00
Subcontractor		\$5,000.00	\$1,000.00
TBD			\$0.00
Total Other Direct Costs		\$11,164.00	\$8,556.00

Detail/calculations on items below per contract budget (to be added)

Direct Costs	\$0.00
Overhead	\$0.00
Total Cost	\$0.00
Fee	\$0.00

Total (Task budget and Task Balance after adding overhead and fee)	\$18,636.00	\$12,876.00
---	--------------------	--------------------

NATIONAL ASSESSMENT GOVERNING BOARD

APPENDIX D: Schedule of Board Quarterly Meetings 2008-2010

Meeting Dates	Location
July 31-August 2, 2008	Mandarin Oriental Washington DC 1330 Maryland Avenue SW Washington DC 20024 (202) 554-8588
November 20-22, 2008	The Westin Arlington Gateway 801 N. Glebe Road Arlington, VA 22203 (703) 717-6200
March 5-7, 2009	Mandarin Oriental Washington DC 1330 Maryland Avenue SW Washington DC 20024 (202) 554-8588
May 14-16, 2009	Hyatt at Olive 8 1635 8 th Avenue Seattle, WA 98101 (206) 774-1234 (phone number will change in 12/08 after opening)
August 6-8, 2009	Mandarin Oriental Washington DC 1330 Maryland Avenue SW Washington DC 20024 (202) 554-8588
November 19-21, 2009	To be determined (possibly Raleigh)
2010 Board Meeting Dates March 4-6, 2010 May 13-15, 2010 August 5-7, 2010 November 18-20, 2010	To be determined

U.S. Department of Education

CONTRACTOR PERFORMANCE INFORMATION

Contractor Name and Address (Identify Division)	1. Contract Number:	
(Please correct the above as needed.)	2. Type of Contract:	
	3. Contract Value (Current plus any unexercised options): \$	
	4. Period of Performance (including any option periods):	
	5. Description of Requirement:	
6. Ratings. Summarize contractor performance and circle or type in the number below that corresponds to the performance rating for each category. Please see the attachment, which explains the rating scale.		
Quality:	0 1 2 3 4	Comments:
Problem Resolution:	0 1 2 3 4	Comments:
Cost Control:	0 1 2 3 4	Comments:
Timeliness:	0 1 2 3 4	Comments:
Business Relations:	0 1 2 3 4	Comments:
Customer Service:	0 1 2 3 4	Comments:
7. Total score:		

Evaluated by:

Agency/Organization _____ Date _____

(In accordance with the Federal Acquisition Streamlining Act, the following information will not be released to the contractor.)

Name and Title:

Telephone Number:

Signature:

E-Mail Address:

Please return this form to the following address:

U.S. Department of Education
National Assessment Governing Board
800 North Capitol Street, NW
Suite 825
Washington, DC 20002-4233

Or e-mail to: Stephen.Swearingen@ed.gov
Or fax to: [202-357-6945](tel:202-357-6945)
Attn: [Stephen Swearingen](#)
RFP # [ED08R0028](#)

SOURCE SELECTION INFORMATION—SEE FAR 3.104

Information entered on this form will be used in source selection decisions and is protected under subsection 3.104 of the Federal Acquisition Regulation. Do not disclose information entered on this form to the contractor or to any other person except as authorized by the Department of Education contracting officer.

Supplementary Questions

To assist the Department of Education contracting officer, we would greatly appreciate your taking the time to answer the following questions, if any, related to the contractor's past performance:

Contractor Performance Evaluation

Instructions for Completing Contractor Performance Information Form

Based on the rating area elements presented below and the rating guidelines on the back of this sheet, please evaluate contractor performance in each of the rating areas. On the “Contractor Performance Information” form, circle (or type in the “Comments:” area) the rating from 0 to 4 that most closely matches your evaluation of the contractor’s performance. Please add written comments for each rating. If you wish, you may attach additional comments or information. We would also appreciate your answers to the specific questions, if any, on the back of the form. Please return the form to the address indicated on the back of the form. Thank you for your time and your cooperation.

The Department of Education will use the information from this form to evaluate offerors competing for contract awards. We may release the information from this form to the contractor during negotiations or debriefings. If we release information from this form, we will not release your name to the contractor.

Elements within Each Rating Area

Quality of Product or Service

- Compliance with contract requirements
- Accuracy of reports
- Appropriateness of personnel
- Technical excellence

Problem Resolution

- Anticipates and avoids or mitigates problems
- Satisfactorily overcomes or resolves problems
- Prompt notification of problems
- Pro-active
- Effective contractor-recommended solutions

Cost Control

- Within budget
- Current, accurate and complete billings
- Costs properly allocated
- Unallowable costs not billed
- Relationship of negotiated costs to actual
- Cost efficiencies

Timeliness of Performance

- Meets interim milestones
- Reliable
- Stays on schedule despite problems
- Responsive to technical direction
- Completes work on time, including wrap-up and contract administration
- No liquidated damages assessed

Business Relations

- Effective management
- Use of performance-based management techniques
- Business-like concern for the customer's interests
- Effective management and selection of subcontractors
- Effective small/small disadvantaged business subcontracting program
- Reasonable/cooperative behavior
- Effective use of technology in management and communication
- Flexible
- Minimal staff turnover
- Maintains high employee morale
- Resolves disagreements without being unnecessarily litigious.

Customer Service

- Understands and embraces service and program goals
- Team approach with the customer
- Satisfaction of end users with the contractor’s service
- Positive customer feedback
- Prompt responses
- Courteous interactions
- Effective escalations and referrals
- Initiative and proactive improvements
- Creative service strategies

Rating Guidelines

Quality of Product or Service

0 - Unsatisfactory	Nonconformance jeopardizes the achievement of contract goals; default.
1 - Poor	Nonconformance requires major agency intervention to ensure achievement of contract goals; show cause or cure notices.
2 - Fair	Quality meets specifications in most cases, however, some agency intervention required to ensure achievement of contract requirements.
3 - Good	Quality meets specifications in all cases.
4 - Excellent	Quality exceeds specifications in some cases.

Problem Resolution

0 - Unsatisfactory	Inadequately resolved problems jeopardize contract goals.
1 - Poor	Significant agency intervention required to resolve problems jeopardizing contract goals.
2 - Fair	Some agency intervention required to resolve problems jeopardizing contract goals.
3 - Good	Successfully overcomes or resolves all problems and achieves contract goals with minimal agency intervention.
4 - Excellent	Anticipates and avoids most problems and successfully overcomes all unforeseen problems.

Cost Control

0 - Unsatisfactory	Cost increases jeopardize achievement of contract goals; or billings routinely include unallowable costs.
1 - Poor	Significant cost increases; or some inaccurate billings including some with unallowable costs.
2 - Fair	Minor cost increases; or some inaccurate billings, but a minimal (1-2) number with unallowable costs.
3 - Good	Contractor performed within costs; but some late billings, none with unallowable costs.
4 - Excellent	Costs were less than the amount cited in the contract; and billings accurate and timely.

Timeliness of Performance

0 - Unsatisfactory	Delays jeopardize the achievement of contract goals.
1 - Poor	Other significant delays.
2 - Fair	Minor delays.
3 - Good	All deliverables on time.
4 - Excellent	All deliverables on time with some ahead of schedule; or stays on schedule despite unforeseen circumstances.

Business Relations

0 - Unsatisfactory	Unethical or illegal business practices.
1 - Poor	Business practices are not attuned to customer support.
2 - Fair	Business practices are somewhat attuned to customer support.
3 - Good	Business practices focus on customer support.
4 - Excellent	Highly effective, proactive business practices focused on customer support.

Customer Service

0 - Unsatisfactory	Response to service requests is routinely late, ineffective, or rude; customers express frustration or anger about many interactions; complaints are unresolved; contractor seems unaware of service issues.
1 - Poor	Response to service requests is often late, ineffective or rude; some complaints are resolved.
2 - Fair	Response to service requests is uneven in timing or effectiveness; customer interactions are tenuous; contractor is trying hard and understands service issues.
3 - Good	Response to service requests is timely, effective and courteous; customers express positive feedback; delivery of service is smooth and organized; collects customer feedback; customer problems are resolved well.
4 - Excellent	Response to service requests is timely, effective and courteous; the contractor is proactive in building good relations with customers, proposing new service strategies, analyzing and reporting on service loads and collecting and using customer feedback.