

Setting Standards for the 1998 NAEP in Civics and Writing: Finalizing the Achievement Levels Descriptions

Susan Cooper Loomis and Patricia L. Hanick
ACT, Inc.

October 2000

Setting Standards for the 1998 NAEP in Civics and Writing: Finalizing the Achievement Levels Descriptions

Susan Cooper Loomis and Patricia L. Hanick
ACT, Inc.

October 2000

The work for this report was conducted by ACT, Inc. under contract ZA97001001 with the National Assessment Governing Board.

Copyright © 2000 by ACT, Inc. All rights reserved.

Table of Contents

| | |
|--|----|
| Executive Summary..... | ii |
| Purpose | 1 |
| Importance of Achievement Levels Descriptions..... | 1 |
| Finalizing the Achievement Levels Descriptions for the 1998 NAEP..... | 2 |
| Step 1: ACT Staff Prepared the Preliminary ALDs for Review by Focus Groups | 2 |
| Writing..... | 2 |
| Civics..... | 3 |
| Step 2: ETS Staff Analyzed the Preliminary ALDs for Writing in Comparison to the Generic Scoring Rubrics and Modified the Generic Scoring Rubrics for Each Type of Writing | 3 |
| Step 3: Focus Groups Were Convened to Recommend Ways to Improve the Preliminary ALDs | 4 |
| Focus Group Participants..... | 4 |
| Focus Group Meetings..... | 5 |
| Focus Group Procedures..... | 6 |
| Focus Group Outcomes | 7 |
| General Recommendations from the Writing Focus Groups:..... | 8 |
| General Recommendations from the Civics Focus Groups: | 8 |
| Step 4: Experts Reviewed the Focus Groups’ Recommendations and Revised the Preliminary ALDs Accordingly | 8 |
| Expert Review Panel Participants..... | 8 |
| Expert Review Panel Meetings..... | 8 |
| Expert Review Panel Procedures | 9 |
| Expert Review Panel Outcomes | 9 |
| General Modifications Made by the Expert Review Panels: | 10 |
| Step 5: Key Individuals Who Had Been Involved in the Development of NAEP Evaluated the Modified ALDs and Recommended Additional Changes | 10 |
| Collecting Opinions from Content Committees | 10 |
| Collecting Opinions from Focus Groups | 11 |
| Step 6: ACT Analyzed the Content Outlined in the Civics Framework in Relation to the Revised Civics ALDs .. | 11 |
| Step 7: ETS Analyzed the Civics Assessment Item Pool in Relation to the Content Described in the Revised Civics ALDs | 12 |
| Step 8: NAGB’s Achievement Levels Committee Authorized the Use of the Modified ALDs for the Pilot Studies and ALS Procedures in Both Civics and Writing | 12 |
| Conclusions of the ALD Finalizing Process..... | 12 |
| Recommendations for Development of Achievement Levels Descriptions | 13 |
| References | 13 |

| | |
|------------|---------------------------------|
| Appendix A | Writing Focus Group Material |
| Appendix B | Civics Focus Group Material |
| Appendix C | Writing Generic Scoring Rubrics |
| Appendix D | Civics Expert Review Material |
| Appendix E | Writing Expert Review Material |
| Appendix F | Civics Follow-Up |
| Appendix G | Writing Follow-Up |

EXECUTIVE SUMMARY

The procedure implemented for finalizing the achievement levels descriptions (ALDs) used in the 1998 NAEP Achievement Levels-Setting (ALS) process in Civics and in Writing is recommended as a model for developing achievement levels descriptions to use for setting standards.

The process of developing achievement levels descriptions has changed somewhat over time. In the 1992 ALS process, panelists wrote the operational achievement levels descriptions at the ALS meeting. For the 1992 ALS process, the ALDs were not finalized until after the ALS panels had adjourned. Beginning in 1994, panelists evaluated the preliminary ALDs and modified them as necessary during the ALS meeting. For the 1998 NAEP, instead of the ALS panelists writing or modifying the preliminary ALDs at the meeting, the descriptions were finalized *before* convening the ALS panels. Participants no longer were responsible for refining the ALDs. This change was an improvement in the ALS process for several reasons:

- It reduced the amount of work time required of panelists in an already full ALS schedule.
- It decreased the cognitive demands placed on panelists during the ALS session.
- It allowed more time to deliberate fully and develop the ALDs than could be made available during the ALS meeting.
- It allowed more stakeholders than just the ALS panelists to be involved in the overall achievement levels-setting process.
- It enabled a rigorous evaluation of the ALDs before they were used in the ALS process.

While the plan to finalize the ALDs was generally judged to be a positive change in the process, ACT, NAGB, and TACSS shared two major concerns about this change prior to implementation of the ALS process. First, it was not clear that panelists would internalize the ALDs sufficiently. Because the ALDs were “givens” to the process, panelists would not have hand-on experiences with developing or

modifying the descriptions. Working with the ALDs had served as the means for earlier panelists to internalize the descriptions. Similarly, since the ALDs were “givens,” it seemed likely that panelists would not “buy in” and identify with the standards to the same extent that had been the case for previous panels. We were concerned that panelists would be less committed to the ALDs and to the standards they set since they had no role in writing the descriptions of what students should know and be able to do. Those concerns appear to have been unfounded, however. Panelists evaluated their understanding of ALDs as positively as had been the case in previous procedures.¹

The process of finalizing the ALDs covered several steps, some of which were iterative. At the recommendation of the Technical Advisory Committee on Standard Setting (TACSS), ACT staff worked to “clean up” the preliminary achievement levels descriptions that were incorporated in the NAEP Framework documents for civics and writing. The writing ALDs, for example, were modified so that a developmental progression across levels and grades was more evident. The civics ALDs were organized into several different formats to facilitate understanding of the great number of highly detailed descriptions.

Focus groups were conducted to evaluate the preliminary ALDs in both writing and civics. Persons were selected to serve on focus group panels in four regions of the nation. Criteria similar to those used in selecting achievement levels-setting panels were used for selecting focus group members. The panels made recommendations for changes to increase the usefulness and reasonableness of the ALDs. Those changes were evaluated by panels of content experts and incorporated into revised achievement levels descriptions.

The revised descriptions were evaluated by a larger group of stakeholders and further, albeit

¹ See, for example, Loomis & Hanick (2000).

relatively minor, modifications were made to the descriptions. After several iterations of review and revision, the ALDs were submitted to the Achievement Levels Committee of the National Assessment Governing Board (NAGB) for approval. The committee approved the ALDs for use in the pilot studies and ALS procedures. NAGB officially adopted the achievement levels—descriptions, cutscores, and exemplar items—during their August 1998 Board meeting.

Setting Standards for the 1998 NAEP in Civics and Writing: Finalizing the Achievement Levels Descriptions

Susan Cooper Loomis and Patricia L. Hanick
ACT, Inc.

PURPOSE

This report is an account of the process used to finalize the statements of what students should know and be able to do in civics and in writing for the 1998 National Assessment of Educational Progress (NAEP). The process was designed to transform the preliminary statements of what students should know and be able to do into the final descriptions used to set the 1998 NAEP achievement levels. The National Assessment Governing Board (NAGB) had already produced assessment frameworks, including preliminary achievement levels descriptions for each subject. A national consensus process was followed to develop those documents.

The assessment frameworks and preliminary achievement levels descriptions were used to develop the item pools and scoring rubrics for civics and the writing tasks and scoring rubrics for writing.

The framework, preliminary achievement levels descriptions, and item/task pools had received extensive and intensive review and scrutiny by many expert review groups before being approved by NAGB.

IMPORTANCE OF ACHIEVEMENT LEVELS DESCRIPTIONS

Congress created the National Assessment Governing Board (NAGB) as the policy-making body to oversee NAEP. Congress also mandated the Governing Board to generate achievement levels that are “useful, reasonable, and valid.”² NAGB has adopted general policy definitions of

student achievement that describe Basic, Proficient, and Advanced performance in broad terms with no reference to a specific subject or grade level. Since 1994, preliminary achievement levels descriptions (ALDs) have been part of the NAEP assessment frameworks. The preliminary ALDs have been the first step in the process to operationalize the NAGB policy definitions of achievement: that is, to specify what students should know and be able to do for each grade level in a specific subject. Because the description of student achievement at different levels of performance is critically important in reporting the outcomes of NAEP, the development of achievement levels descriptions is a vital part of this mandate.

In past NAEP achievement levels-setting (ALS) procedures, the job of developing the ALDs was one of three major tasks carried out by panelists participating in the standard setting sessions. Their responsibilities included:

- Developing content-based descriptions of each level of achievement for each of the three grades assessed by NAEP;
- Determining numerical cutpoints on the assessment that tied the achievement descriptions to performance on NAEP; and
- Selecting assessment items that illustrated the skills and knowledge characterizing student performance at each of the three achievement levels at each of the three grades.

In the 1992 ALS process, panelists wrote operational achievement levels descriptions at the ALS meeting. Beginning in 1994, panelists evaluated the preliminary ALDs and modified them as necessary during the ALS meeting. For the 1998 NAEP, instead of the ALS panelists writing or modifying the preliminary ALDs at

² Public Law 103-382, *Improving America's Schools Act of 1994*.

the meeting, the descriptions were finalized *before* convening the ALS panels. Participants no longer were responsible for refining the ALDs. This change was an improvement in the ALS process for several reasons:

- It reduced the amount of work time required of panelists in an already full ALS schedule.
- It decreased the cognitive demands placed on panelists during the ALS session.
- It allowed more time to deliberate fully and develop the ALDs than could be made available during the ALS meeting.
- It allowed more stakeholders than just the ALS panelists to be involved in the overall achievement levels-setting process.
- It enabled a rigorous evaluation of the ALDs before they were used in the ALS process.

FINALIZING THE ACHIEVEMENT LEVELS DESCRIPTIONS FOR THE 1998 NAEP

NAGB’s approach to setting achievement levels is based on collecting opinions from a broadly representative sample of knowledgeable constituents who are invited to become involved in the ALS process. In keeping with NAGB’s consensus-building approach, ACT designed a method to collect opinions through a series of focus groups, expert reviews, and discussions of the preliminary ALDs. The preliminary ALDs were modified and refined according to the recommendations collected during the review process prior to convening the ALS panels. The method used to develop the ALDs involved several steps.

The focus groups represented a broad segment of the population. They evaluated the preliminary ALDs to determine whether the ALDs were both useful and reasonable. The judgment of “reasonableness” was with regard to the NAGB policy definitions of Basic, Proficient, and Advanced achievement. Content experts in civics and writing reviewed the results of the focus groups’ evaluation and modified the ALDs accordingly. The revised ALDs were then reviewed by several different groups of informed stakeholders, including the original focus group members. As a final review, the content described in the revised Civics ALDs

was compared with the content described in the *Civics Framework* and the Civics NAEP item pool. After extensive review and evaluation, ACT recommended the revised ALDs to NAGB for approval for the Civics NAEP and the Writing NAEP. A more detailed explanation follows of the method that ACT implemented for developing the ALDs for the 1998 NAEP.

Step 1: ACT Staff Prepared the Preliminary ALDs for Review by Focus Groups

The preliminary ALDs for both civics and writing received a great deal of attention before the focus groups reviewed them. In preparation for the focus group meetings, ACT analyzed the achievement levels descriptions in their preliminary form. The preliminary ALDs appeared as lists of bulleted statements in each framework. To assist focus groups in understanding the knowledge and skills described in the ALDs, ACT created tables using the bulleted statements. The tables displayed the descriptive statements for each grade across achievement levels, and for each achievement level across grades.

During the process of preparing the preliminary ALDs for the focus groups, ACT identified several idiosyncrasies in both the civics and the writing preliminary descriptions. ACT consulted the Technical Advisory Committee for Standard Setting (TACSS) who recommended that the preliminary ALDs be “cleaned up” before the focus groups were convened. The process involved many iterations of modifications and review for both the descriptions and table formats.

Writing

One major issue in the preliminary writing ALDs was the terminology used as “anchor words” for the description. The key terms for one achievement level description often were replicated in the description for a different achievement level. For example, the adjective “appropriate” was a key word in the description of writing at the Basic level. However, “appropriate” also appeared as a key descriptor for Proficient and Advanced writing. TACSS advised ACT to revise the statements so that

different achievement level descriptions would not share identical “anchor words.”

Further, identical statements were used to describe writing skills associated with student performance within a specific level of achievement across different grades. For example, one phrase of a preliminary ALD stated that students in grade 4 performing at the Basic level should be able to “demonstrate [an] appropriate response to the task in form, content, and language.” Exactly the same phrase was used to describe performance of students in grade 12 at the Basic level. TACSS advised ACT to modify the preliminary ALDs so that the statements would reflect developmental skill progression across grades at the three achievement levels. This change represented a departure from the plan by the framework developers to have the developmental progression addressed in the difficulty of writing tasks and the demands of the scoring rubrics. Nonetheless, recognizing that a goal of the ALDs is to increase the ability of the public to interpret accurately the reports of student performance on NAEP, this modification appeared necessary. It seemed unlikely that the public would feel comfortable with a set of achievement levels describing performance for twelfth grade students performing at the Basic level in exactly the same terms as used to describe performance at the Basic level for fourth grade students.

There was also concern that the quality of writing described for Basic level performance was higher than “partial mastery,” which was the NAGB policy definition of Basic performance. Related to this issue was the alignment of the generic scoring guides to the achievement levels descriptions, which is described in greater detail in Step 2 below.

The ACT NAEP ALS Project staff worked extensively with ACT staff who had served on the project team for developing the 1998 Writing NAEP Item Specifications, revising the framework, and developing the preliminary achievement levels descriptions. They helped modify the preliminary achievement levels descriptions and develop the charts, tables, and

other content-related materials for focus group meetings. Please see Appendix A for the materials developed for the writing focus group meetings.

Civics

The preliminary ALDs for civics were very precise statements that resembled actual assessment items. The civics preliminary ALDs had been written in specific language to guide the item development process. However, the long lists of detailed statements seemed to present too great a challenge for meaningful evaluation by focus group members and standard-setting panel members. ACT created tables that grouped the statements according to the five areas outlined in the *Civics Framework* as the “Organizing Questions and Content.” A member of the civics framework and item development panel and a NAGB Staff member responsible for framework development evaluated the initial organization of the tables and recommended several modifications, which ACT implemented. Many versions of the tables were reviewed and revised. In the end, the civics focus group members worked with three different table formats that displayed the preliminary ALDs across achievement levels by grade, across grades by achievement level, and by content area across achievement levels for each grade. Please see Appendix B to review the materials developed for the civics focus group meetings.

Step 2: ETS Staff Analyzed the Preliminary ALDs for Writing in Comparison to the Generic Scoring Rubrics and Modified the Generic Scoring Rubrics for Each Type of Writing

In preparation for the focus group meetings, ACT compared the preliminary ALDs with the generic scoring rubrics for the Writing NAEP. ACT concluded that the scoring guides for responses scored as 1, 2, and 3 described writing that most likely would be classified as Below Basic, while the scoring guides for responses scored as 4, 5, and 6 described writing that most closely matched the descriptions for Basic, Proficient, and Advanced, respectively. The generic scoring rubrics are in Appendix C. ACT

was concerned that the achievement levels would be set too high to be considered reasonable and useful if a score of 4 out of a possible 6 were required for Basic performance. Further, several key adjectives directly and clearly linked the scoring rubrics to the preliminary ALDs. For example, the scoring rubric for a response scored as 4 stated that the writing exhibited “word choice (that) was basic.” The language of the scoring rubric inadvertently linked the score of 4 to the Basic achievement level category, leading ALS panelists to associate a specific scoring rubric value with a specific achievement level. The Technical Advisory Committee on Standard Setting (TACSS) advised ACT to resolve these issues before using the ALDs for focus group meetings. ACT expressed these concerns to Educational Testing Service (ETS) who revised the scoring guides accordingly without disturbing the calibration of the prompts.

Step 3: Focus Groups Were Convened to Recommend Ways to Improve the Preliminary ALDs

Focus Group Participants

Abt Associates Inc., under contract to ACT, conducted one focus group in each of the four NAEP regions for writing and one in each region for civics. Abt used a networking process to identify nominators of candidates to participate in the focus groups held in Massachusetts, Georgia, Missouri, and Washington. Abt mailed information explaining the project to state assessment and testing directors in the four NAEP regions. Abt then followed with a telephone call to the directors to request the names of candidates and colleagues who could nominate qualified focus group participants within the region. For example, in the Northeast, Abt staff contacted assessment and testing directors in Massachusetts, New Hampshire, Connecticut, and Rhode Island. The directors themselves would nominate candidates and often recommended other nominators who were familiar with curriculum experts and outstanding teachers in local school districts or throughout the state.

Abt also contacted professors who taught incoming freshmen in English composition or U.S. government, and educators who were involved in curriculum development and teacher education at the college level. In addition, Abt spoke with education editors and journalists for major newspapers in the host cities and major employers, charities, and civic groups like the League of Women Voters and the Bar Association.

Abt identified between 25 and 30 nominators for each city. The number of nominees ranged from 67 in St. Louis to 116 in Atlanta. Due to the December holidays, it was particularly difficult to recruit nominees in the St. Louis area. In the Portland region, candidates were reluctant to participate because of professional commitments that involved reorganizing the local school districts and dealing with recently announced cuts in school funding. In spite of these obstacles, however, Abt recruited an outstanding group of participants in each region.

Panels consisted of a mix of teachers, nonteacher educators, and general public members with 70-75% representing educators and 25-30% representing the general public. In order to assure proportional representation on the panels, Abt used the following guidelines to select the panel members:

- Six classroom teachers (two per grade)
- Three or four nonteacher educators (not more than two for any grade)
- One to three members of the general public (not more than one for any grade)
- Geographic variation within each NAEP region
- Equal numbers of each gender
- Racial/ethnic representation reflecting national proportions

Focus group participants were familiar with the knowledge and skills in either writing or civics that students typically possess in at least one of the grades tested by NAEP. In addition, focus group members were trained in the subject tested by NAEP, or had experienced in work-related areas. When recruiting teachers, Abt considered the number of years of teaching experience and

the grade levels taught, awards for excellence in teaching, and involvement in developing state or district standards, assessments, or curricula. Teachers came from private, parochial, and public schools. Nonteacher educators were primarily curriculum experts in the subject tested by NAEP. Representatives of the general public were involved with education in some way, such as working with students on writing skills, or participating in programs that exposed student to law, government, and concepts of civil society.

The number of focus group participants ranged from 8 to 12. A total of five recruited persons failed to participate in the four civics focus groups (1 in Boston, 2 in St. Louis, 2 in Portland), and 3 persons were absent from the writing focus groups (2 in Boston, 1 in Portland). One-third of the participants (28) were males and two-thirds were females (57). The civics groups were evenly split by gender, but the writing groups were predominantly females (84 %). Minorities included African American, Native American, and Hispanic participants. A full list of participants for each writing focus group is included in Appendix A and for each civics focus group in Appendix B. Also included in these appendices are tables displaying the composition of the panels and examples of several professional biographies of panelists.

Focus Group Meetings

Abt convened the panels on Saturdays according to the following schedule: Boston on November 15, 1997; Atlanta on December 7, 1997; St. Louis on December 14, 1997; and Portland, Oregon on January 10, 1998. Participants received mileage reimbursement and free parking. Some individuals who traveled greater distances than could be reached reasonably by car received airfare and hotel accommodations. Refreshments were provided.

Each meeting took place in a large hotel conference room equipped with a screen and overhead projector. Morning and afternoon sessions alternated for each subject such that two focus groups for each subject were held in the morning, and two were held in the afternoon.

The writing focus groups lasted three to three and one-half hours, and the civics groups lasted four hours because of the complexity of the materials they were required to review. The morning session started at 9:00 and the afternoon session started at either 1:30 or 2:00.

Observers included representatives of NAGB and NAGB staff, members of the Expert Review Panel, Abt Associates staff, and ACT staff. An experienced focus group facilitator from Abt Associates led the meeting. Abt staff took notes during the discussion, tape-recorded the discussion, and used a laptop computer and LCD panel to display the group's work during the discussion. Abt used these materials to prepare reports of the general and specific comments for the Expert Review Panels.

Two weeks before the meeting, focus group members received extensive advance materials to review prior to the meeting. The information assisted them in understanding the ALDs. Please see Appendix A (writing) and Appendix B (civics) for samples of the packets of advance materials for each subject, which include the following:

- A brochure on the National Assessment Governing Board with the general policy definitions for Basic, Proficient and Advanced performance;
- The appropriate *Writing* or *Civics Framework* for the 1998 National Assessment of Educational Progress;
- Tables showing individual statements of the preliminary achievement levels descriptions arranged in cells by achievement levels within grade, and by grades within achievement level;
- For civics, a third table formatted statements for the three achievement levels at one grade, ordered by the five content questions outlined in the *Civics Framework*;
- A pamphlet, "Multiple Challenges," which gives background information about the 1998 NAEP assessments;
- A cover letter with instructions for focus group members to follow to prepare for the meeting;
- A map and directions to the meeting.

The cover letter informed focus group participants of their upcoming task: to make recommendations about the preliminary ALDs to an Expert Review Panel who would incorporate the information into final recommendations to NAGB. Panelists were instructed to read and develop an understanding of the NAGB policy definitions of Basic, Proficient, and Advanced performance. They were directed to read the framework, to learn how the assessment was developed, and to study the preliminary achievement levels descriptions.

The advance packets also contained tables that showed single statements of the preliminary achievement levels descriptions arranged in cells by achievement levels within grade, and by grades within achievement level. The cells illustrated the interrelationships among the statements of achievement at the three levels and at the three grades. While there were many possible ways of ordering these statements, the significance of the tables was to assist participants in understanding the performance requirements at each grade and achievement level.

The statements for civics were formatted three different ways. The first was a single page of statements grouped by content that described one level of achievement for one grade. The second format was statements for three levels of achievement at one grade ordered by the five content questions outlined in the *Civics Framework*. The third iteration was the same as the second without the content question headings. Examples of these ALD tables are included in the writing and civics advance packets in Appendix A and B. During the course of the project, ACT slightly modified the format of the original ALD tables by assigning a reference number to each cell. This addition greatly facilitated panelists' identification of statements during discussions.

Focus Group Procedures

The overall approach to facilitating the focus group meetings centered on members discussing carefully crafted questions about various qualities of the achievement levels descriptions.

After a general group discussion of these guided questions, the members made specific recommendations that were intended to improve each description, at each grade, for each achievement level. The recommendations were summarized and reviewed again by the entire group. The group made final corrections and the meeting was adjourned. The complete process was documented.

The meeting started with the Abt Project Director welcoming the members of the focus groups and introducing NAGB, ACT and Abt staff, and any members of the Expert Review Panel who were attending. A NAGB staff member then welcomed the participants, described the function of NAGB, and presented background information about NAEP. The ACT Project Director followed, and she described the relationship between the NAGB policy definitions and the ALDs. She also explained the role of the ALDs in the NAEP standard setting process.

At this point the meeting was turned over to the focus group facilitator who set the rules for the discussion. Participants introduced themselves and briefly described their involvement with students in the subject area. The facilitator followed the *Focus Group Discussion Guide* for all of the focus group meetings. Please see Appendix A to review the guides for writing and Appendix B for civics. She began by explaining the purpose of the meeting and instructing the panelists in the reference materials. She lead the group in discussing three main questions:

- Given the NAGB policy definitions of the three achievement levels (solid academic performance, partial mastery, and superior performance) are the descriptions *reasonable* statements of what students should know and be able to do at each grade?
- Are the statements *useful* in describing student performance? That is, are they effective in communicating student performance?
- Given what you know about students, are the achievement level descriptions *reasonable* statements of what students

should know and be able to do at each NAEP grade level?

For the first group in Boston, the order of the main questions was slightly different. The discussion began with the question about the reasonableness of the achievement level descriptions (question 3 above). The Boston group discussed Proficient performance at each grade followed by Basic and ending with Advanced. The Boston group then reviewed each performance level across all grades. And finally, the Boston group reviewed the usefulness of the descriptions as a means to report student performance. Staff concluded that the discussion guides needed to be changed because the format caused group discussion to be constrained and cut short on several occasions.

Based on the Boston experience, Abt modified the discussion guides to allow more time for general discussion followed by a period during which panelists made recommendations that related to specific ALD statements. This modified approach appeared to work well for the writing focus groups, but not as well for the civics groups. Because the preliminary ALDs for civics had been written as long lists of very specific statements, civics group members had difficulty distinguishing them from test items. They were inclined to discuss the ALDs as if they were test questions. Members tended to make recommendations to restate specific descriptions rather than to make recommendations to combine and summarize the descriptions. After two sessions, ACT decided to change the task somewhat for the civics focus groups.

The procedures previously outlined for facilitating the civics focus groups were followed for the last two meetings in St. Louis and Portland, with one exception. Members of the civics focus groups were asked to recommend general, broad, “umbrella-type” statements for the ALDs. To do this, members worked in small, grade groups to generate the recommendations based on the following questions:

- Is this description essential for solid performance? (Partial mastery? Superior performance?)
- If this description were omitted, would the sense of solid performance be lost? (Partial mastery? Superior performance?)
- Which elements can be combined to make the statements more useful?

At the end of each focus group, members completed an evaluation questionnaire about the implementation of the meeting. Panelists gave their opinions about the materials, facilities, and group discussions. They also had the opportunity to write open-ended comments about their experience as a focus group member. The results of these questionnaires and comments are included in Appendix A for writing and Appendix B for civics.

Focus Group Outcomes

The focus group members took their responsibilities seriously and commented at length on many of the individual descriptions. In general the participants found their task to be daunting. Even when instructed otherwise, they continued to think of the descriptions as test items or scoring guides. They often expressed a desire to see the test items and were frustrated by their lack of access to this information. ACT staff and NAGB staff as well as the Expert Review Panel informed the focus group participants of their specific role in the overall achievement levels-setting process. Staff encouraged participants to work within the constraints that had been set for this process, which the group found difficult to follow.

Overall, the eight focus groups produced an abundance of rich, qualitative information about the preliminary ALDs. The change in the civics group discussion format produced many more suggestions than were received from the first two civics groups. The modified format seemed to help members produce recommendations that were more relevant to the work of the Civics Expert Review Panel. A complete set of recommendations was compiled for each focus group for each city. These comments have been collected in Appendix A for writing and Appendix B for civics as part of the advance

mailings to the Expert Review Panels. The following is a brief summary of the general recommendations made by the focus groups.

General Recommendations from the Writing Focus Groups:

- The ALD statements in general described achievement that was too high for first draft writing.
- The description of Basic achievement in particular was too high for 4th and 8th grades, but was about right for 12th grade.
- The description of Advanced achievement for 12th grade should be higher.
- The statements should describe a clearer progression of writing skill development across levels and grades.
- The descriptions should be stated in simpler, clearer language.
- The descriptions were too vague and should be more specific and use examples.
- There seemed to be a gap between the achievement described for 8th grade and 12th grade.

General Recommendations from the Civics Focus Groups:

- The description of Basic achievement was too high for all grades because it denoted more than “partial mastery.”
- The content described as Basic achievement was too difficult, especially at the 4th grade.
- The description of Advanced achievement was truly “superior performance” for 4th and 8th grades.
- The statements should describe a clearer progression of civics skill development across levels and grades.
- The descriptions should be stated in simpler, clearer language. The language of the civics preliminary ALDs was inflated and most statements were too complex.
- The descriptions of achievement for 4th grade should be based on concrete experiences that involve home, school, and community. The statements for 4th grade describe concepts that were too abstract to be developmentally appropriate for 4th grade achievement.

- The level of detail was inconsistent in describing achievement at the different grades. In particular, the descriptions for 12th grade achievement were very broad, whereas for 8th grade they were very specific.

Step 4: Experts Reviewed the Focus Groups’ Recommendations and Revised the Preliminary ALDs Accordingly

Expert Review Panel Participants

Within a few weeks of the last of the focus group meetings, the Expert Review Panels (ERP) for Civics and Writing met to review the recommendations from the focus groups. The Expert Review Panels included members from the NAEP Framework Panels, the NAEP Item Development Committees, and the NAEP Standing Committees for each content area. The Writing ERP consisted of five experts, three of whom attended the ERP meeting, and the Civics ERP consisted of eight experts, seven of whom attended. All but one ERP member participated in a focus group meeting, and at least one ERP member was present at each focus group session. Please see Appendix D for a complete list of the Civics Expert Review Panel members and their affiliations, and Appendix E for the Writing Expert Review Panel.

Expert Review Panel Meetings

The ACT Project Director conducted the ERP meetings. Observers included a representative of NAGB staff, Abt Associates staff, ACT staff, and an editing consultant for ACT. Please see Appendix D for a complete list of civics attendees and Appendix E for writing attendees. The entire process was documented. Abt staff took notes during the discussion, tape-recorded the discussion, and used a laptop computer and LCD panel to display the group’s work during the meeting.

The ERP meetings were held on weekends in Abt Associates’ Chicago offices at the end of January 1998. They lasted two days for writing and three days for civics. Members received advance materials that included a summary report of the comments and recommendations generated by all of the focus groups in addition

to all of the materials utilized during the focus group sessions. Those materials consisted of the focus group discussion guide, a list of the focus group participants and their credentials, a set of achievement level descriptions, and a description of the logistics for the focus group meetings.

Expert Review Panel Procedures

The overall approach to the ERP meeting involved a thorough review of the preliminary ALDs and the focus group recommendations. The ERP then revised the ALDs through an iterative process with numerous drafts produced throughout the meeting. The ERP worked to produce a final version of the ALDs that was agreeable to all members. The meeting format included the following procedures:

- The ERP reviewed and discussed the preliminary ALDs and the recommendations offered by the focus groups. The review was *very* thorough; every recommendation was considered for each ALD.
- After lengthy group discussion, the ERP reached preliminary agreement on how to revise, edit, modify, and reassign each statement.
- With the assistance of an editor, the ERP revised the statements and drafted a narrative version of the ALDs. The narrative was in both a summary version and an extended, complete version for each ALD.
- Additional polishing of the narrative version of the descriptions took place following the ERP meeting. The polished version was distributed to the ERP for further consideration. After minor adjustments, members agreed to approve the revised descriptions as the version of the ALDs they recommended to NAGB.

Working together as a group, the writing ERP agreed on how to modify the ALDs and completed their work in bulleted format. Each person then drafted a narrative version of the ALDs for one of the three grade levels. They discussed this version and reached agreement on the general format that would serve as the “prototype” for the finished product. They also agreed on the descriptions that would serve as

the summary paragraphs. They permitted the editor to polish the final descriptions, subject to their approval.

The civics ERP spent many hours reviewing each description with respect to the recommendations from the focus groups. At the end of the second day, they had reached a preliminary agreement on how to revise each ALD in bulleted format. The editor worked overnight to incorporate those statements into narrative statements and summary paragraphs. On the third day, the ERP identified areas in the narrative statements that needed additional work, and they modified those working in smaller, grade groups. They then reconvened as a single group to review and modify the ALDs. At the close of the meeting, the revised paragraphs and summary statements were given to the editor for polishing, subject to ERP approval.

Expert Review Panel Outcomes

The process of reviewing the preliminary ALDs was thorough and the revisions were extensive. The ERP made every effort to strike a balance between implementing all of the recommendations offered by the focus groups, and maintaining a cohesive progression of skill development across the grades. In general, ERP members addressed the specific concerns expressed by the focus groups. In nearly all instances, the revisions followed the recommendations from the focus groups. In a few cases, however, the panel agreed that the recommendations were inappropriate and those were not implemented. They revised the ALDs to be unambiguous statements that used fewer jargon-type terms. The group was attentive to providing adequate coverage for the subject area, as outlined in the frameworks. The Civics ERP systematically addressed each content domain outlined in the *Civics Framework*. Members considered skill progression across all three grades to assure that the level was developmentally appropriate. Because there were so few members of the writing panel, they worked as a single group, for the most part. While they initially felt that the ALDs would require separate statements for each type of writing (narrative, informative, and persuasive),

they determined during the review process that this would not be necessary. One additional important decision reached by this group was that no description of *voice* would be included in the final version of the descriptions, although *voice* had been described in the preliminary writing ALDs.

A complete set of modifications was compiled for each of the expert review panels. Full documentation of the process is found in Appendix D for the civics review and Appendix E for the writing review. The following is a brief summary of the general changes made to the descriptions by the ERP.

General Modifications Made by the Expert Review Panels:

- Adjusted the degree of difficulty described in the statements of achievement according to the recommendations of focus groups
- Revised the language in the statements to be clear and simple
- Used examples to make the descriptions specific
- Combined separate statements that described related ideas
- Clarified the distinction between achievement levels
- Revised the statements to describe a clear developmental progression of skills across levels and grades
- Revised the statements so that the level of detail in the descriptions was consistent across levels and grades

Step 5: Key Individuals Who Had Been Involved in the Development of NAEP Evaluated the Modified ALDs and Recommended Additional Changes

TACSS reviewed the process used by ACT to develop the descriptions of the achievement levels for the 1998 Writing and Civics NAEP. On March 26, 1998 TACSS agreed unanimously that the process had been reasonable and that the final version of the ALDs appeared to provide descriptions that were consistent with NAGB policy statements and frameworks. TACSS

recommended that ACT take three additional actions:

- Have the NAEP Framework Planning Committee members review the final version of the ALDs;
- Have the Civics NAEP item pool reviewed with regard to the final version of the civics ALDs; and
- Make a complete, detailed account of the finalizing process available to the larger community.

Collecting Opinions from Content Committees

Once the recommended versions of the ALDs were finalized, ACT requested comments about the recommended ALDs from the original focus group members and key people who had been involved in the development of the NAEP for writing and civics. Members of the NAEP Standing Committee, the NAEP Item Development Committee, and the NAEP Framework Planning Committee for each subject were asked to review the modified ALDs. ACT contacted 40 civics participants and 15 writing participants. Each was mailed copies of the recommended ALDs, the assessment framework, and the generic NAGB policy definitions of Basic, Proficient, and Advanced performance. ACT asked the content committee members to review the materials and make comments and suggestions for improvement to the recommended descriptions. A copy of the surveys and the results are in Appendix F for civics and Appendix G for writing.

Comments were received from 26 persons in civics and 7 in writing. Persons who served on the content committees generally approved of the modified ALDs.

- When asked if the revised ALDs should be adopted without substantive changes, 19 (73.1%) of the civics members and 5 (71.4%) of the writing members said “yes.” To review these data please see Tables F-1 and F-2 in Appendix F for civics and Tables G-1 and G-2 in Appendix G for writing.
- Of those who recommended changes, one or two members thought the descriptions were too high for some grades and levels, while

one or two other members thought they were too low.

- The greatest number of recommendations was made for the 4th grade Advanced level in civics. Four members thought that description was too demanding.
- Three of the 7 writing members sensed a “gap” in the progression of skill development described by the grade 8 Basic and Proficient writing ALDs. They thought the Basic ALD described writing that seemed too low, and the Proficient ALD described writing that seemed too high at the 8th grade.

Collecting Opinions from Focus Groups

ACT asked the original focus group members to review the modified ALDs. Eighty-five participants were mailed copies of the revised ALDs and the generic NAGB policy definitions of Basic, Proficient, and Advanced performance. The focus group members were informed that a telephone interviewer would call them to ask their opinions of the recommended descriptions. Their opinion was to be based on the reasonableness and usefulness of the revised descriptions, relative to the NAGB policy definitions. A toll-free telephone number was provided to members for returning calls to ACT. A copy of the telephone survey and the results are in Appendix F for civics and Appendix G for writing.

The results of the telephone survey indicated that the persons who served on the focus groups generally approved of the modified ALDs.

- When asked if the revised ALDs seemed reasonable, 89.2% of the civics members and 84.6% of the writing members responded positively.
- When asked if the revised ALDs were clear and easily understood, 78.4% of the civics members and 79.5% of the writing members replied “yes.”
- When asked if the revised ALDs reflect the three achievement levels, between 94.6% - 100% of civics members agreed (depending upon the level), and between 84.6% - 92.3% writing members agreed.

When asked if additional changes to the ALDs were necessary, 40.5% of the civics members and 48.7% of the writing members replied “yes.” Comments from these individuals consisted mainly of editorial type suggestions for the statements. Very few made recommendations for changes that would likely lead to significant alterations of the levels. Although the percentages of participants who advocated additional changes seemed rather high, 78.4% of civics members and 82.1% of writing members recommended that the ALDs be adopted exactly as they appeared in the survey.

The writing survey included one question that did not appear in the civics survey. This question (#7) asked participants to evaluate the descriptions for 8th grade Basic and Proficient, relative to one another. Specifically, members were asked whether or not there appeared to be a “gap” in the progression of skill development described by the two levels. Three members of the writing content committees surveyed earlier thought a gap existed. Of the 29 focus group respondents to this question, 10 (34.5%) said that there was a “gap:” that is, Basic achievement seemed too low and Proficient too high for grade eight.

Step 6: ACT Analyzed the Content Outlined in the Civics Framework in Relation to the Revised Civics ALDs

The preliminary ALDs for both writing and civics were changed substantially as a result of the finalizing process. It was important to ascertain that the revised descriptions were not changed to such an extent that they no longer reflected the content described in the assessment frameworks. ACT analyzed the civics content areas described in the *Civics Framework* as they related to the content described in the revised civics ALDs. The revised ALDs, as defined in the framework, described all five areas of civic knowledge. More specifically, the ALDs described the civics content outlined in the framework for all three grade levels, for all three achievement levels, and for a broad range of intellectual skills.

Step 7: ETS Analyzed the Civics Assessment Item Pool in Relation to the Content Described in the Revised Civics ALDs

At the request of NAGB, Educational Testing Service further analyzed the revised ALDs by classifying the items in the civics exercise pool according to the revised descriptions. Sixty-five different descriptive statements were examined across the three grades. There were only two sentences – one at 4th grade Basic and the other at 8th grade Basic – that had no assessment exercises associated with them. The findings of the ETS analysis are summarized below:

- The overall match between the content of the civics item pool and the content of the revised achievement level descriptors was strong.
- Some descriptive statements could be associated with four or five assessment items, while others could be associated with twenty items.
- A limited number of items from the exercise pool could not be directly associated with a descriptor.

At each grade, over 90% of the civics NAEP exercise pool could be classified according to the revised achievement level descriptions. The number of exercises that could not be classified was small. At grade four, the content of 6 of the 90 items could not be classified as described by the ALDs. At grade eight, 10 of the 151 items could not be classified and at grade twelve, 7 of the 152 items could not be classified. The majority of the unclassified exercises in grades 8 and 12 were designed to measure students' abilities to interpret information presented in tables, graphs, or charts. This skill had been included in the assessment specifications, but had not been described in the ALDs. After consulting members of the ERP, ACT added a description at the grade 8 Proficient level and grade 12 Basic level to require skills in reading and interpreting maps, charts, and other tabular and graphical data. The other items that could not be classified measured students' abilities in different content areas that had not been described in the ALDs. The ERP members judged the items to be non-significant in content, so no additional descriptors were added.

Step 8: NAGB's Achievement Levels Committee Authorized the Use of the Modified ALDs for the Pilot Studies and ALS Procedures in Both Civics and Writing

After the steps above had been completed and approved by TACSS, ACT reported the final results to the NAGB Achievement Levels Committee and asked for approval of the ALDs in writing and civics. During the August 1998 NAGB meeting, the Achievement Levels Committee gave provisional approval of the ALDs for use in the pilot study and the operational ALS process. The committee was unwilling to make the ALDs the final, official achievement levels descriptions for the subjects until after the achievement levels had been finally set. If the results of the ALS procedure were approved by NAGB, then the ALDs would be part of the results officially approved at that time. If not, then other ALDs would potentially be needed to describe the performance of students at the levels decided upon by NAGB. The committee judged that it was premature to officially adopt the ALDs prior to having all of the ALS results.

CONCLUSIONS OF THE ALD FINALIZING PROCESS

The 1998 achievement levels descriptions received a great deal of attention, considerably more than had been given to ALDs in previous NAEP achievement levels-setting processes. The method of collecting opinions regarding the ALDs was thorough and diligent. Great effort was given to incorporating the collected opinions into the recommended ALDs.

The ALDs were further polished as a result of the continuing analyses that compared the content of the revised ALDs with the content of the item pool. Descriptors were added to the 8th and 12th grade civics ALDs stating that students should be able to interpret information presented in tables, graphs, and charts. At the suggestion of writing ALS panelists, ACT added an explanation of testing conditions as a preamble to the writing ALDs. And finally, the word "draft" was removed from the writing ALDs and

replaced with statements indicating that students should produce finished work “within the time allowed,” or “given the time constraint.”

The Achievement Levels Committee of NAGB met on April 29, 1998 to review the process ACT used to develop and refine the recommended ALDs, and to formulate their recommendations to NAGB regarding the ALDs. The Committee was satisfied with the format for the ALDs that included a shorter summary paragraph in italics followed by a longer, more detailed paragraph describing performance. The Committee recommended that NAGB authorize the use of the civics and writing ALDs in the 1998 process for developing the NAEP achievement levels. At the NAGB meeting in August 1998, the Board authorized ACT to use the revised achievement levels descriptions for the NAEP pilot studies and the ALS meetings. NAGB adopted the achievement levels in May 1999.

RECOMMENDATIONS FOR DEVELOPMENT OF ACHIEVEMENT LEVELS DESCRIPTIONS

The role of the ALDs is of central and primary importance in the NAGB/ACT ALS Process. Assuring that the ALDs are reasonable and useful is essential to the successful production of valid achievement levels. ACT recommends that a process similar to that described in this report be used to develop ALDs, and that this procedure be completed before frameworks are published and widely distributed.

Given the need to have the framework and achievement levels descriptions guide the development of the assessment items and scoring rubrics, the ALDs should be finalized rather early in the process. ACT recommends that the framework development committees continue to develop the preliminary ALDs, but that more time and emphasis be given to this part of their work. The ALDs must be subjected to careful review and evaluation by broadly representative panels of stakeholders. As was the case in the process implemented by ACT, the process should be iterative so that broad-based input can be shared with experts and the work of

the experts can then be shared again with broadly representative samples of stakeholders.

Achievement levels descriptions should guide development of items and scoring rubrics for assessments. There should be a careful review of the correspondence between the item pools and achievement levels descriptions *before* the assessment item pools are finalized. The purpose of this comparative review is to assure that all—or most—items can be matched conceptually to an ALD and, more importantly, to assure that there are no ALDs for which there are no items. This comparison should be conducted by a panel of sufficient size and expertise to provide some assurance of the reliability of the outcome. This comparison must be made with the full understanding that the judgments reached by this evaluation may not correspond to the empirical results based on student performances. That is, students performing in the range of the Basic level of achievement will not necessarily perform well on items that were judged to match or represent the Basic achievement level description. There should be a relatively high correspondence between the two, but there will not be a one-to-one correspondence.

Taking these steps to finalizing the ALDs prior to setting the achievement levels should help to assure a successful ALS process resulting in achievement levels that are judged to be useful, reasonable, and valid.

REFERENCES

Loomis, S.C. & Hanick, P.L. (2000). *Setting achievement levels for the 1998 NAEP in Civics: Final report*. Iowa City, IA: ACT.