

A Response to "Setting Reasonable and Useful Performance
Standards" in the National Academy of Sciences'
Grading the Nation's Report Card^{1,2}

Ronald K. Hambleton
University of Massachusetts Amherst

Robert L. Brennan
University of Iowa

William Brown
Brownstar, Inc.

Barbara Dodd
University of Texas at Austin

Robert A. Forsyth
University of Iowa

William A. Mehrens
Michigan State University

Jeff Nellhaus
Massachusetts Department of Education

Mark Reckase
Michigan State University

Douglas Rindone
Connecticut Department of Education

Wim J. van der Linden
University of Twente, The Netherlands

Rebecca Zwick
University of California, Santa Barbara

May 10, 1999

¹The authors serve as technical advisors to ACT on the NAEP standard-setting process.

²The authors are grateful to Susan Loomis of ACT, Inc., and Mary Lyn Bourque, of the National Assessment Governing Board, for providing technical assistance needed in the preparation of this paper.

Our Position on the NAEP Performance Standards

Since 1990 the National Assessment Governing Board (NAGB) has been setting performance standards (called “achievement levels” by NAGB) on the National Assessment of Educational Progress (NAEP). Performance standards have been set on the 1990 and 1992 Mathematics, 1992 Reading, 1994 US History, 1994 Geography, 1996 Science, 1998 Civics, and 1998 Writing assessments. The purposes of the performance standards are to (1) provide a frame of reference for policy-makers, educators, and the public to understand the NAEP test results; (2) provide more interpretive information about the meaning of NAEP test scores by carefully defining three categories of performance: Basic, Proficient, and Advanced; and (3) promote excellence in education. Standard-setting on NAEP has been controversial—initially, critics were opposed to both performance standards and the process used for setting them. Today, performance standards appear to be widely accepted and valued, but the process for setting them remains controversial.

Our review of the evidence for the NAEP performance standards indicates there is considerable support for using the current methods for setting performance standards on NAEP and other national and state assessments. The NAGB standard-setting methods themselves have been carefully and systematically developed and improved through numerous research studies, pilot-tests, and implementations over eight years. Every aspect of the process is state-of-the-art in standard-setting methodology, from the selection of broadly representative groups of panelists and advanced preparation for the standard-setting meetings, to the development of performance descriptors, to the training of panelists and implementation of the methods in a five-day process, to the careful evaluation and documentation of the process.

Our positive evaluation of the NAEP standard-setting methodology is clearly at odds with the very negative evaluation of this methodology reported by the National Academy of Sciences (NAS) in Chapter 5 of **Grading the Nation's Report Card** (Pellegrino, Jones, & Mitchell, 1999). "Chapter 5: Setting Reasonable and Useful Performance Standards," in our view, presents a very one-sided and incomplete evaluation that is based largely on dated and second-hand evidence. The faulty conclusions and incomplete review of evidence to support the conclusions suggest to us that the authors of Chapter 5 (1) placed too much emphasis on dated material, (2) ignored the extensive body of research that supports the validity of the NAEP performance standards, and (3) were relatively uninformed about the challenges associated with the NAEP standard-setting process. In our judgment, Chapter 5 of the NAS Report does not conform to generally accepted scientific standards of objectivity in conducting research or evaluation. In sum, we are surprised, dismayed, and disappointed with the inadequate scholarship reflected in the review of NAGB's eight years of research and development to set performance standards on NAEP. Policy-makers, educators, and the public deserve a much better review and evaluation than they received for their substantial investment of research funds. It should be noted that at least some of our concerns were well-known to the NAS prior to the release of its Report. For example, Robert L. Brennan sent a letter to the NAS in October, 1998, after he reviewed a draft of Chapter 5. In that letter he stated that the chapter "makes strong statements based on a series of reports and derivative documents..., but the Report fails to make reference to other reports and papers that seriously challenge the data, logic, and the conclusions that are cited in Chapter 5." He went on to say that such omissions "make Chapter 5 not credible as a scholarly research or evaluative document" and he called upon the NAS "to see fit to retract Chapter 5 in its current form and issue a revised version..." The NAS response to Brennan was that, "The

NRC [National Research Council] and the committee stand by the report and its conclusions and recommendations.” Their only modification of the draft text was the inclusion of less than two lines of text. Such a minimal change is hardly responsive to the substance or the spirit of Brennan’s letter, or to the more detailed concerns raised in this response to Chapter 5 of the NAS Report.

Our Response to Chapter 5 of the NAS Report

The primary purpose of this paper is to respond to the criticisms of the NAEP standard-setting methodology put forth in Chapter 5. Specifically, the objectives of this response are as follows:

1. Identify several points made in Chapter 5 with which we agree.
2. Offer several general criticisms of Chapter 5.
3. Address the specific criticisms regarding the NAEP standard-setting process noted in Chapter 5.
4. Consider the evolution and integrity of the achievement levels-setting process that has been used with NAEP assessments.

In the next section, the goals and conclusions of Chapter 5 are summarized. Subsequent sections address the four objectives noted above.

Goals and Conclusions of the NAS Report

According to the authors of Chapter 5, their goals were the following (p. 163):

1. To provide an overview of NAEP’s performance standards and the achievement levels-setting process as it was conducted through 1996.
2. To summarize the major findings of previous evaluations and research efforts that have examined this process.

3. To present a detailed accounting and evaluation of the achievement levels-setting process as it was applied to the 1996 NAEP Science Assessment.
4. To present the committee's conclusions and recommendations about standard-setting.

These authors arrived at the following summary conclusion: "Standards-based reporting is intended to be useful in communicating student results, but the current process for setting NAEP achievement levels is fundamentally flawed" (p. 162). In the remainder of this paper, we will consider this summary conclusion and the evidence that is offered in support of it.

We note, however, that the phrase "fundamentally flawed," used throughout the NAS Report as an overall evaluative conclusion, is never justified in any serious scientific sense. This mantra has certainly captured the attention of the public, the press, and the Congress. Rhetoric should **not** be substituted for scientific inquiry, and continued repetition of an unwarranted conclusion does not enhance its validity.

Points of Agreement

There are a number of observations and comments in Chapter 5 with which we concur. The most important of these are discussed below.

We definitely agree that standards-based reporting "has the potential to be a significant improvement in communicating about student achievement to the public and to policy makers" (p. 171). Most state departments of education also agree, and currently use standards-based reporting in their own assessments. Furthermore, we strongly support the call for additional research to (1) help improve the use and interpretation of standards-based reporting, (2) assess the impact of standards-based reporting on curriculum development and instruction, and (3) continue to explore the usefulness of alternative standard-setting procedures. In fact, as noted

later in this paper, a considerable amount of research related to alternative standard-setting procedures has been undertaken by NAGB and more is being planned.

We also support the recommendation in Chapter 5 that preliminary achievement levels descriptions “should guide the development of assessment items and exercises” (p. 177). In fact, in recent assessments (e.g., Civics) it is our understanding that such descriptions have played an important role in the development of the assessment.

We concur with the authors of Chapter 5 that all standard-setting methods rely heavily on informed judgments and that “true achievement levels do not exist” (p. 173). We were surprised, however, to see the authors of Chapter 5 offer this as one of their conclusions. In fact, this view has been in the measurement literature for more than 20 years, and is accepted by every knowledgeable measurement specialist known to us who works in the area of standard setting.

In addition, we support the conclusion that despite the judgmental nature of standard-setting, the process “whereby decisions are reached in achievement levels-setting should be well documented and clearly communicated” (p. 173). In a later section of this paper, we argue that considerable documentation exists for the standard-settings that have been undertaken for the NAEP assessments. The NAS evaluation reveals no evidence to suggest reference to available documentation for the 1996 Science achievement levels-setting process.

General Criticisms

Before reacting to the more substantive issues in Chapter 5, it seems appropriate to make a few observations about the general style of argumentation and writing relative to the mission of the NAS. Under Executive Order number 12832 (January 14, 1993), President Bush cited among the functions of the National Research Council the following:

1. To survey the broad possibilities of science, to formulate comprehensive projects of research, and to develop effective means of utilizing the scientific and technical resources of the country for dealing with such projects.
2. To promote cooperation in research, at home and abroad, in order to secure concentration of effort, minimize duplication, and stimulate progress; but in all cooperative undertakings to give encouragement to individual initiative, as fundamentally important to the advancement of science.
3. To gather and collate scientific and technical information, at home and abroad, in cooperation with governmental and other agencies, and to disseminate such information to duly accredited persons and the public.

Adherence to these functions should have resulted in a report based on such generally accepted scientific principles as a balanced and complete referencing of the literature, a careful consultation of primary sources of information, an explicit flow of argument based on clear premises, and a judicious evaluation without unwarranted prejudices. Instead, we believe Chapter 5 represents the results of an evaluation study in which each of these principles is painfully ignored. This conclusion is substantiated by our observations below.

Standard setting has been a prolific area of research in educational measurement.

Measurement journals such as **Applied Measurement in Education**, **Applied Psychological Measurement**, **Educational Measurement: Issues and Practice**, and the **Journal of Educational Measurement** have published hundreds of articles reporting research on standard-setting methods. At national conferences, an even larger number of papers on standard setting have been presented. In 1994, NAGB and NCES organized a joint conference with 19 keynote addresses that together offered a comprehensive review of the state of the art in standard setting

for large-scale assessments (Bourque, 1995). The vast majority of these papers evaluate the Angoff method, propose a modification of the method, or research a variation in its implementation. The authors of Chapter 5 acknowledge that the Angoff method is the “most prevalent approach to standard setting currently in use” (p. 165). Nevertheless, Chapter 5 references only a handful of journal articles, presentations, and chapters in conference proceedings from this massive body of literature.

It should be noted that four of the references in Chapter 5 were added only after a report on the preliminary draft of Chapter 5 was made available to the Technical Advisory Committee on Standard-Setting (TACSS), a committee established by ACT, as required by NAGB, to advise ACT on the standard-setting process. In the discussion about Chapter 5 with the authors, the TACSS voiced strong objections to the lack of balance. After this meeting, the following sentence (15 words only) was added to Chapter 5: “Counter arguments are presented by Hambleton and Bourque (1991), Kane (1994, 1995), Mehrens (1995), and Brennan (1998).”

Two kinds of references are particularly conspicuous by their absence from the list of references. First, since their initiation of standard setting on NAEP, the staffs of NAGB and its contractors have conducted numerous studies, field-testing many variations and improvements of their standard-setting methods. The nature of these studies is reviewed later in our response. A recent list of related studies includes 58 archival publications, 55 technical reports, and 43 conference presentations. Only a small number of the studies are referenced in the NAS Report. In particular, given the extensive criticisms in the chapter on the 1996 standard-setting process, we missed references to “Setting Achievement Levels on the 1996 National Assessment of Educational Progress in Science: Final Report” (ACT, 1997a, 1997b, 1997c, 1997d, 1997e). The five volumes of this report, together over 1200 pages of documentation, describe all procedures,

technical decisions, research projects, pilot studies, and results and include all meeting agendas, briefing books, and training material for the setting of achievement levels on the 1996 NAEP Science Assessment. We know that all of these documents were provided by NAGB to the authors of Chapter 5 for their review. Nevertheless, it seems the authors chose to ignore or exclude these and other primary sources in preparing their review.

Second, an important source of information for a professional review is publications in scientific journals. This category is important because the process of independent, blind review through which these journals select their articles embodies the generally recognized principle that critique by peers is an integral part of the process of ensuring scientific quality. Chapter 5 lists no more than ten titles from this category. Instead, the authors chose to rely almost exclusively on an earlier evaluation of NAEP and the Trial State Assessment by the National Academy of Education (NAE). Nearly half of the references in Chapter 5 are to these National Academy of Education reports or accompanying publications. All of this work was carried out on the 1990 and 1992 standard-setting initiatives. Rather than a scholarly, independent evaluation based on up-to-date primary information and publications that have withstood the critique of fellow-researchers, the NAS researchers seem to have chosen to appeal to the authority of NAE, the General Accounting Office (GAO), and a small group of evaluators with well-known convictions about standards and the Angoff method.

Not only are the references in Chapter 5 inadequate, the chapter is also weak with respect to other aspects of the scientific approach. For example, such an approach would begin with a careful delineation of the object of evaluation, a description of the approach, and an explanation of the evolution over nearly a decade in the methods of standard setting used by NAGB. It

would then reach conclusions through a clear chain of logic. Instead, after a handful of introductory pages, the authors summarize a few “selected findings from past NAEP evaluations and research” (p. 166), and then immediately jump to the conclusion that the “achievement levels are flawed” (p. 167).

In addition, the authors have failed to recognize even such a simple thing as the three qualitatively different stages in the NAGB achievement levels-setting procedure. The procedure starts with a description of the achievement levels for the subject area, which are based on NAGB’s policy definitions of the basic, proficient, and advanced levels. In the next stage, panelists are trained and asked to translate the achievement level descriptions into performance standards on the NAEP score scales. In the last stage, NAGB reviews the previous work and adopts or modifies the recommended achievement levels.

Because Chapter 5 treats all stages as one big stage, it is difficult to associate Chapter 5 conclusions with specific aspects of the standard-setting process. Consider, for example, the 1996 Science achievement levels-setting process, which resulted in a very small percent of students in the advanced category. If this result is viewed as problematic was it due to a flaw in the version of the Angoff method used to translate the level descriptions into achievement levels? Or, did the method perform well, but was there a flaw in the procedure used to establish the achievement levels descriptions? Or, were the science achievement levels descriptions just too challenging for students? Or, was the NAEP science item pool more difficult than pools for other standard-settings?

The same lack of logic is present in the casual conclusion that the standard-setting process is flawed because of its sensitivity to item response format (p. 166). Is the procedure responsible for this presumed “flaw”? Or, are there persistent differences in student motivation? Or,

differences in item scoring? Should the explanation lie in the fact that the achievement level descriptions are more in line with the scoring rubric of some types of items? Or, would there possibly have been problems in scaling the data? Should our conclusion be that there was no problem at all but that the results are a reflection of the way judges react to test items? In short, what arguments do the authors have to support their immediate conclusion that the standard-setting method was flawed? There appear to be many possible explanations for the results, and it is not obvious that flaws in standard-setting are necessarily the explanation.

We believe that Chapter 5 was based on an approach that is lacking in scientific credibility. As already indicated, Chapter 5 seems at variance with important principles related to the functions of the NAS. Likewise, Chapter 5 does not meet important utility, feasibility, and accuracy standards formulated in the program evaluation standards published by the Joint Committee on Standards for Educational Evaluation (1994). For example, consider Standard A-10, "the conclusions reached in an evaluation should be explicitly justified, so that stakeholders can assess them." This standard and many of the other standards were clearly violated in the NAS evaluation of NAGB's achievement levels-setting process.

Specific Criticisms

The authors of Chapter 5 concluded that the current standard-setting procedures used with the NAEP assessments are fundamentally flawed. This conclusion was based primarily on three factors. A brief discussion of each of these factors follows.

Believability of the Results

The first reason cited in Chapter 5 in support of the conclusion concerns the believability of the results, particularly the small percentages of students judged to be advanced in science. To

support this claim, the authors note that NAGB revised the science advanced achievement level recommended by the achievement levels-setting panelists, with the result that a larger proportion of students was classified as advanced than would otherwise have been the case. Moreover, the authors claim that numerous external comparison studies conducted by the NAE (e.g., Burstein, Koretz, Linn, Sugrue, Novak, & Baker, 1996; McLaughlin et al., 1993; Shepard, Glaser, Linn, Bohrnstedt, 1993) support the conclusion that NAEP proficient and advanced cutscores are consistently set too high.

However, we have serious concerns about the NAE panel's cursory analysis of data and evidence that leads them to the general conclusion that the NAEP achievement levels place too few students in the advanced levels. The following observations are offered to illustrate our concerns:

1. Based on the occurrence of small proportions of students in some categories (e.g., science advanced), the NAE panel appears to draw the inference that the process of arriving at achievement levels is flawed. We do not concur that the small proportions are evidence that achievement levels are flawed. High achievement levels (which lead to small proportions) are an expected consequence of rigorous achievement level descriptions. We strongly believe, for example, that the advanced achievement levels descriptions in science are indeed challenging. High achievement levels and small proportions of student scores above them are both expected and consistent with the descriptions. Consequently, it seems incongruous to criticize the achievement levels-setting process (i.e., the process of collecting ratings that lead to the achievement levels) as resulting in achievement levels that are too high.

2. The logic proposed by the NAS panel seems to be that “the current procedures are fundamentally flawed” (p. 174), because, in part, “the results are not believable,” based on the claim that “too few students are judged to be advanced” relative to other conceptions of advanced. This logic is puzzling on three counts. First, it suggests that someone knows, before the fact, something about what the answer (i.e., the standards) should be, or what the answer should not be. If the Chapter 5 authors truly believe that the answer can be ascertained, then they should inform NAGB, so NAGB can decide as a matter of policy whether or not to incorporate such preconceptions into the standard-setting process.

The second puzzling aspect of the authors’ logic and assertions is the presumption that results from testing programs unrelated to NAEP are an obvious and appropriate benchmark for evaluating NAEP results. Such comparisons are, at best, problematic when the testing programs differ substantially from NAEP and/or have different purposes.

The third puzzling aspect of this logic is the apparent assumption that the word “advanced” means (or should mean) the same thing in different testing programs. In particular, there is no reason to think that the word “Advanced” in the Advanced Placement program has any particular bearing on the word “Advanced” as a description of one of the achievement levels in NAEP.

3. The authors also fail to acknowledge that an extraordinarily large number of the individuals participating on the various NAEP standard-setting panels have indicated satisfaction with the process and confidence in the results of the process. Since 1992, NAGB has set over 60 achievement levels. The achievement levels range across seven content areas and three grade levels. Within grades and subjects, the achievement levels

resulted from the work of standard-setting panels constituted by approximately 20 to 30 individuals. Throughout each of the standard-setting sessions, panelists were asked to complete extensive questionnaires designed to ascertain their satisfaction with the process and confidence in the results. In almost all cases the panelists indicated strong satisfaction with the process and confidence in the results. There were only three instances in which the results of the standard-setting process were challenged (i.e., the three achievement levels for Grade 8 Science, 1996), and in those instances, by only a small number of the panelists.

4. The Chapter 5 authors point to the lack of correspondence between NAEP achievement level results and external evidence of student achievement in other contexts (e.g., Advanced Placement examinations/courses) to support their assertion that too few students attain the advanced level on NAEP. There are sensible uses of Advanced Placement data in ongoing validation efforts related to the achievement levels, although Advanced Placement data do not provide an unambiguously clear criterion for validating any specific NAEP achievement level. In particular, the content of the Advanced Placement examinations and the testing conditions are very different from the NAEP. Even so, as noted below, there is some evidence that the percentage of students scoring at high levels on the Advanced Placement science test corresponds reasonably well with the percentage of students at the advanced level on the Grade 12 Science NAEP.

In order to provide a context for judging the recommendations for the levels on the 1996 NAEP Science Assessment, NAGB requested information on the percentage of grade 12 public school students who would have been eligible for selection into the NAEP sample and who did receive a grade of 3, 4, or 5 on any AP Science test in high school.

The College Board database was scanned for the 1994 sophomores, 1995 juniors, or 1996 seniors who met the criteria (without duplication). The final results indicated that 1.93% of the grade 12 cohort took an AP science course and received a score of 3, 4, or 5 on an AP high school science test.

5. Linn (1999) reported, for 11 states, the comparisons between the percent of proficient and advanced students identified by NAEP and the state assessment. For two of the states, students actually appeared to perform better using the NAEP achievement levels than the states' own standards. For four of the other states, the results were relatively close. These findings certainly do not send out shock waves about the invalidity of NAGB's achievement levels because they were set too high. The achievement levels may be high, but Linn's findings do not provide evidence of invalidity. On the contrary, the findings may raise more concerns about the performance standards in states where the differences are very large. For example, in one of the states, NAEP results using the achievement levels suggested that about 15% of the students were proficient and advanced. Using the states' own performance standards, the figure was 90%!

Item Types and Item Difficulty

The second reason given by the authors of Chapter 5 to support the conclusion that NAEP's standard-setting procedures and results are fundamentally flawed is that "the achievement levels-setting results vary significantly depending on the type and difficulty of the items used in the judgment process" (p. 174). Specifically, constructed response items (e.g., short response, extended response, and performance items) typically result in higher standards than

those set using multiple-choice items, and the standards obtained using easy items are different from those obtained using difficult items.

In a review of the early NAE studies that the NAS claims support this conclusion, Kane (1993, 1995) argued, we think persuasively, that such evidence was not necessarily an indictment of the Angoff method. Although the authors of these NAE studies viewed the differences in results based on different item types or item difficulty to be method artifacts that should have been eliminated by the panelists, Kane observed that this was not the only reasonable interpretation.

He states:

One could assume that the apparent difference in the quality of student performance between dichotomous items and extended response items . . . are real, and that students are meeting judges' expectations, on recognition tasks, but not doing as well, relative to judges' expectations, on tasks that require an extended response. . . . The fact is that many scholars believe extended response items tap aspects of student achievement not directly assessed by multiple-choice items. (Kane, 1995, p. 125)

If, in fact, the two types of items are measuring different aspects of student achievement, then different levels of achievement may be expected.

Brennan makes a similar point:

Some find this [i.e., different performance standards for dichotomous items versus extended response items] to be reasonable, even expected, based on the fact that the two types of items are intended to measure different content and/or constructs. (Brennan, 1998, p. 9)

Brennan further noted that if such differences are to be used as evidence of the invalidity of the standard-setting procedure, then it must be shown "that the scaling procedures used in NAEP should lead to similar results for both types of items and that the lack of congruence is attributable to inadequacies in the standard-setting and not inadequacies in the scaling" (p. 9).

Kane (1995) also raises the scaling issue with respect to the interpretation of differences between the achievement levels set on the basis of either hard or easy items. He notes that to attribute such differences to the standard-setting method ignores the possibility that

the NAEP scaling system, which is probably the most complicated such system in use, [may not be] doing a good enough job in adjusting for item difficulty to ensure that estimates of scores on the IRT-based achievement scale based on the hardest items would generally be equivalent to the estimates based on the easiest items. (p. 127)

Kane continues by noting that:

Research on this type of vertical scaling has not generally been very encouraging, and, therefore, it would seem prudent to at least consider the possibility that the scaling procedure might be having an impact... (Kane, 1995, p. 127)

Another factor that must be considered in this discussion is the differential motivation of examinees with respect to the two types of items. It is well documented, for example, that the proportion of students omitting extended response items is greater than the proportion omitting multiple-choice items. Likewise, DeMars (1998) also observed that for low-stakes tests students were more likely to omit extended response items than multiple-choice items.

Difficulty in Estimating Probabilities

The third reason Chapter 5 authors use to support their claim that the procedures are flawed is that panelists have difficulty estimating probabilities for test items (p. 175). These authors agree with a previously made claim that the judgment task posed to raters "is too difficult and confusing" to categorize performance at "three different levels" (p. 166).

There are several problems with this claim, as pointed out in a variety of earlier arguments (e.g. Cizek, 1993; Kane, 1995; Zieky, 1995). First, as recognized by virtually all researchers, including the NAE panel, the Angoff procedure is the most prevalent approach to standard setting

in use (p. 165). To conclude, based on two references (see p. 175) that panelists cannot do something that panelists have been doing for decades seems inappropriate. It is certainly contrary to evidence collected from NAEP standard-setting panels.

This is not to suggest that standard setting is easy. Making judgments about levels of performance that correspond to advanced, proficient, and basic will be difficult no matter what specific procedure is implemented. Certainly there is evidence in the literature that making probability estimates is difficult. However, the particular cognitive decision being made in the standard-setting process is one that panelists have been thoroughly trained to do. The standard-setting process is a five-day process that involves thorough training and iteration of tasks with feedback to the panelists between the iterations. The design and implementation of the process is largely aimed at helping the panelists with their tasks, and there is considerable evidence that panelists can perform their tasks very well. For example, as discussed later in this paper, in each grade of every standard-setting process, two groups of panelists work independently. Comparisons of these mini-replications do not suggest that panelists are behaving irrationally or erratically. On the contrary, these comparisons strongly suggest that the panelists are quite consistent in their final judgments.

In discussing the cognitive complexity of the task, it should be mentioned that there is legitimate debate about just how panelists interpret their task. Zieky, for example, suggests the following interpretation:

I don't have any research to verify this, but my experience in standard setting leads me to believe that judges involved in the major item-judgment methods of setting standards are not really making estimates of the performance of some hypothetical group of minimally competent examinees.

I think the actual process is closer to one in which a judge looks at an item and decides that he or she would not be willing to call an examinee minimally competent unless the examinee had a chance of at least x of getting the item right. I believe that judges are not making sloppy estimates of probability. They are

directly expressing their own values about what level of performance they would consider good enough to call minimally competent. (Zieky, 1995, p. 30)

The accuracy of the NAE claim that the Angoff method requires an impossible cognitive task was directly addressed by Cizek as follows:

Specifically, the central claim that the Angoff methodology presents an "impossible cognitive task" rests on speculation that the provision of three Angoff ratings...for a single item is an overly taxing departure from the usual Angoff procedure in which only one rating is generated. The single study that bears directly on this issue was commissioned by the NAE and is reported in McLaughlin (1993a). The clearly stated finding of this study—not highlighted in the NAE report—is presented below:

"The major finding from this study is that the use of the three-ratings-per-item did not introduce substantial and statistically significant artifactual regularity into the cutpoints...Therefore, it does not seem warranted, based on these results, to recommend against using the three-ratings-per item methodology." (Cizek, 1993, p. 128)

Based on his analyses, Cizek concluded that:

Put simply, there is no evidence in the psychometric literature to support the Report's contention that the Angoff method requires an "impossible cognitive task." Indeed, the literature presents precisely the opposite conclusion ... (Cizek, 1993, p. 10)

Kane also offered rebuttals to several studies related to the Chapter 5 authors' conclusion regarding the difficulty of the judgments and concluded as follows:

The evidence developed in the field studies of the technical properties of the 1992 NAEP standard setting do not seem to justify the conclusion (Shepard et al., 1993, p. 77) based largely on these studies,

"that the Angoff procedure is fundamentally flawed because it depends on cognitive judgments that are virtually impossible to make." (Kane, 1995, p. 129)

The authors of Chapter 5 also suggest that panelists cannot judge the relative difficulty of items. However, correlations between totally independent Angoff ratings and item difficulties are substantial (see, for example, Hambleton & Bourque, 1991). Further, item judgments are

averaged for each panelist, and the final performance standards are averaged over panelists. The achievement levels are not dependent on the judgment of a single item by a single judge. Rather, for example, setting standards on a 50-item test with a panel of 15 judges would produce 750 item ratings to be averaged for each round of ratings. The NAS authors' concern about the ratings of single items is both unwarranted and misleading. It would be comparable to making an important decision about examinees based on their performance on a single item.

In determining whether a task can be done, one reasonable source of information is the people who have been asked to do the task. Kane addressed this issue as follows:

Nevertheless, the judges asked to complete Angoff ratings managed to do so without complaint in both NAGB standard-setting efforts and the NAE replication (McLaughlin, 1993b) of one of the NAGB studies. The results for NAE's replication were very close to the original NAGB results, and both of these sets of results were in general agreement with the results of NAE's whole book ratings (McLaughlin, et al., 1993). Furthermore, the Angoff method has been used to set passing scores on a host of licensure and certification tests, as well as on numerous state testing programs, without major complaints from the judges involved. Not only can judges do the task, but they also do not seem to find it particularly anxiety provoking as such (although they do sometimes find it tedious). (Kane, 1995, p. 24)

Thus, the conclusion of the authors that "the judgement tasks are difficult and confusing" (p. 182) seems exaggerated and simply not true.

At various places in Chapter 5 (most notably in the last full paragraph on p. 175) statements are made that suggest the authors believe that those involved in NAEP standard-setting have not considered alternative methods. This is not true. Many alternative procedures have been considered, and several have been pilot tested. For example, Kane (1993, 1995) provided some compelling arguments against the contrasting groups method that has been proposed by some of the critics of current approaches. Also, the bookmark procedure has been considered, but it (and other item mapping procedures) requires the arbitrary selection of a

response probability (e.g., 50%, 65%, 74%) that may substantially affect results. In addition, holistic procedures have been pilot tested, but those procedures generally have yielded higher, not lower, achievement levels. Additional details about some of NAEP's exploration of alternative procedures are provided in the next section.

The Integrity of the Achievement Levels-Setting Process

Definition of the Process

The full process for determining the standards that are used by NAGB to report the results of NAEP begins with NAGB's decision to implement standards-based reporting and their development of policy definitions for the achievement levels. The process ends with NAGB formally implementing the use of standards for reporting. There is a certain ambiguity in the characterization of this process in Chapter 5. The breadth or scope of the achievement levels-setting process is not acknowledged, and the technical and policy components of the process are not differentiated. The beginning and end points of the process are policy concerns. The middle, starting with the policy definitions and continuing to developing numerical values on the NAEP score scale, is the focus of this section.

Prior to the science achievement levels-setting, the second (middle) stage of the process included the development of achievement levels descriptions. Currently, preliminary achievement levels descriptions are created as part of the NAEP framework development process, and they serve as a beginning point for the formal achievement levels-setting. In this section, the term "achievement levels-setting" will refer to the process of determining numerical values on the NAEP scale and identifying exemplar items that are consistent with the descriptions. These results are sent to NAGB for policy approval. The end of the full process consists of NAGB's

consideration of both the achievement levels descriptions and the numeric values that define the achievement level boundaries.

Care in Implementation

As noted previously, standards are based on judgments, but they need to be well-considered judgments made by qualified people using a process that is carefully designed to provide the people (judges) with the information they need to formulate or develop their judgments. The process used for setting achievement levels for the NAEP has been designed to meet all of these criteria. The judges involved in the process have been selected because they are well qualified, the process is the result of a continuing program of research and refinement, and the standard-setting process is conducted under the thorough review of a technical advisory committee composed of highly experienced persons in the fields of psychometric theory, educational policy, and standard setting.

The current process for developing achievement levels for NAEP, as specified in a series of documents from ACT, Inc. (ACT, 1997f, 1997g, 1997h, 1998, and others, listed in Appendix A of this paper), is the result of a continuing evolutionary process that is focused on refining the methodology for standard setting. This work began in 1992³ (ACT, 1993b) with the setting of achievement levels in mathematics, reading, and writing and has continued through work on geography, U.S. history, and science, to the current work on civics and renewed efforts to set achievement levels in writing. The current effort related to achievement levels in writing demonstrates the care that has gone into the process. The TACSS did not believe that the original achievement levels for writing (1992 NAEP) met appropriate standards of quality and recommended to NAGB that the standards not be used for reporting (correspondence between

³ACT's work on setting the achievement levels continued earlier work by Hambleton and Bourque (1991).

the ACT project director and NAGB, dated June 14, 1993). NAGB accepted that recommendation. New achievement levels for the 1998 NAEP writing are being produced at this time. Thus, the process is not a blind implementation of fixed methodology. Rather, it is a carefully reviewed and refined process. The results of that process are only put forward for use after careful review by experts in the field.

With each implementation of the achievement levels-setting process, and with each pilot study and research investigation, the process has been refined to address issues that have been raised by ACT, TACSS, NAGB, and by various evaluations. Thus, the process that was implemented for geography and U.S. history in 1994 and for science in 1996 was a refinement of the 1992 process that was evaluated by the NAE (Glaser, Linn, & Bohrnstedt, 1993). These refinements addressed many of the points discussed in the NAE report. It is amazing that Chapter 5 references the 1993 NAE evaluation as a description of the science achievement levels-setting process since the process did not even begin to be developed until 1996! The science process was changed noticeably from the 1992 design evaluated by the NAE, in part because of the large number of performance items. Further, whole booklet feedback was implemented to provide holistic information about student performance to the judges. The use of the prior NAE evaluations as the basis for criticizing the current process is inappropriate because these evaluations do not take into account the adaptations that were made in response to the NAE evaluation and for the structure of the NAEP science assessment.

The achievement levels-setting process has been a very public one. From the very beginning of ACT's contract to develop achievement levels, the process has been open to review. In fact, it is probably the most carefully considered and open standard-setting process in the country, if not the world. The openness of the process, and the critical evaluations that have

taken place, have served as a stimulus for extended research on standard-setting processes. As documented in ACT (1995), many alternative standard-setting procedures have been investigated to determine if they might lead to improvements over the current methods on the basis of replicability of results, psychometric soundness, and reasonableness of standards. For example, as noted previously, whole booklet classification methods, student performance classifications, item mapping, and item response string generation techniques have been studied as alternatives to the current methodologies. In all cases, these methods failed to meet stringent criteria for psychometric soundness and reasonableness of standards, so they were not implemented as replacements for current procedures. It is not that ACT and NAGB have ignored the recommendations of organizations who have evaluated the process; the alternative procedures have been found wanting on a number of criteria.

Each achievement levels-setting process is informed by an array of computer simulation, field test, and pilot test studies. These studies are used to evaluate alternative standard-setting procedures, feedback mechanisms, and training approaches. Additional studies follow the achievement levels-setting providing information about the inferences that are made from the achievement levels. For example, two pilot studies were conducted prior to the operational achievement levels-setting for science. These studies focused on the procedures needed for the hands-on science tasks, research using several response probability criteria in item mapping studies, and the effect of consequences information on the ratings. The studies were also used to check the training procedures and the scheduling of the five-day achievement levels-setting process. Reckase and Bay (1998) provide a summary of the full scope of the achievement levels-setting efforts as well as the many studies of alternative standard-setting procedures that have been performed to inform the achievement levels-setting process.

Achievement Levels-Setting Process

The development of the achievement levels-setting process begins with a technical proposal that is carefully reviewed by NAGB (ACT, 1997a, 1997b). This proposal provides a review of potential standard-setting procedures as well as the recommended approach and plans for a series of research studies designed to address specific issues that have been raised by previous evaluations. For example, specific studies were suggested in ACT (1997a) to investigate the issue of "cognitive complexity" and the impact of consequences data on achievement levels. These proposed studies were reviewed by NAGB. The result is a series of specific planning and design documents (ACT, 1997b, 1997c, 1998) that lay out a program of research and development focused on refining the method to be used for the operational standard setting. The TACSS also reviews the plans.

It should be noted that the panelists consist of teachers at the specific grade level, non-teacher educators, and members of the general public with relevant background and/or experience in the subject matter tested. The use of a broadly representative group of panelists is not only a matter of policy specified by NAGB; it is also a legislative requirement. It is not a choice made by ACT or the TACSS. In making this observation, we are not suggesting that NAGB policy or Congressional authorization are inappropriate, although certain recommendations made by the authors of Chapter 5 (see Appendix D referenced in the chapter) seem to represent at least a partial challenge to both the law and NAGB policy. If the Chapter 5 authors disagree with NAGB's policy, then they should provide a compelling argument with supporting evidence for an alternative policy. Especially, the authors should not confuse such policy or statutory matters with the technical issues involved in arriving at recommended achievement levels. As reflected in

ACT reports of the various standard-setting sessions throughout this decade, the three types of panelists tend to arrive at quite similar recommendations.

The process of selection of these individuals is replicable. Nominators are identified through a stratified random sampling process, insuring both that the panelists represent all regions and specified demographic attributes within the country, and that multiple samples of panelists can be obtained for use in research and pilot studies. The methods used to identify and select qualified panelists are well documented in ACT reports (e.g., ACT, 1997b).

The panelists are well trained and this training has become more extensive as the process has evolved. For mathematics, panelists went through about seven hours of training concerning NAEP, NAGB policy, the achievement levels, and the standard-setting process (ACT, 1993a). The current process begins training with extensive briefing materials that are sent in advance of the session, plus two days of training about NAEP, NAEP scoring, NAGB policy, the standard-setting process and the feedback that panelists receive (ACT, 1998). The type of training has also been refined, now using multimedia presentations and carefully planned instructional tasks.

A unique feature of the achievement levels-setting process is that the process itself involves a mini-replication so that the amount of instability in the standards can be estimated. The achievement levels are determined for two groups working independently in each grade within each achievement levels-setting session. The similarity of the achievement levels set by these two groups has been uniformly high for all of the achievement levels-setting projects. The pilot testing processes have also yielded partial replications with different groups and at different times. These replications also show that the achievement levels that are set are very stable.

The end result of this design and development work is a process that is carefully designed and refined through a series of pilot studies and field trials. It is carefully reviewed at every step in development and implementation. Chapter 5 does not reflect the integrity of this process.

Concluding Statement

Chapter 5 and Appendix D present a disconcerting juxtaposition of perspectives. Chapter 5 aims at characterizing NAGB's achievement levels-setting process as "fundamentally flawed" with hardly a reference to, not to mention a discussion of, numerous reports and papers that document the process and challenge this rhetoric. Basically, Chapter 5 devotes two lines to mentioning alternative perspectives. Then, Appendix D devotes almost six pages to suggesting a very complicated process that has never been subjected to even a single pilot study. It is trivially easy to criticize an extant process if one ignores documentation and evidence, and it is equally easy to assign potential and virtue to a new process if it is not subjected to a reality test. Both perspectives, and particularly their juxtaposition, seem to us to be misinformed, misleading, and basically unsound from a scientific perspective. In particular, the "fundamentally flawed" rhetoric that pervades "Grading the Nation's Report Card" is not just unwarranted—it is simply wrong. No standard-setting process is, or could be, perfect, but the ACT/NAGB process has much to recommend it.

In this paper, we have provided a detailed response to many aspects of the review/evaluation of NAGB's initiatives to set achievement levels on the NAEP. Our review of the evidence suggests that Chapter 5 on "Setting Reasonable and Useful Performance Standards" for NAEP constitutes a one-sided, incomplete and inaccurate accounting of the NAGB/ACT standard-settings conducted during this decade. As such, Chapter 5 is a disservice to NAGB,

educational policy makers, educators, and the public. In our opinion, in this instance, the NAS review and screening process clearly failed to produce a trustworthy and credible evaluation.

References

- ACT, Inc. (1993a). *Description of mathematics achievement levels-setting process and proposed achievement level descriptions: volume I*. Iowa City, IA: ACT, Inc.
- ACT, Inc. (1993b). *Setting achievement levels on the 1992 National Assessment of Educational Progress in mathematics, reading, and writing: a technical report on reliability and validity*. Iowa City, IA: ACT, Inc.
- ACT, Inc. (1995, December). *Research studies on the achievement levels set for the 1994 NAEP in geography and U.S. History*. Iowa City, IA: ACT, Inc.
- ACT, Inc. (1997a). *Setting achievement levels on the 1996 National Assessment of Educational Progress in Science: final report volume I, pilot study 1*. Iowa City, IA: ACT, Inc.
- ACT, Inc. (1997b). *Setting achievement levels on the 1996 National Assessment of Educational Progress in Science: final report volume II, pilot study 2*. Iowa City, IA: ACT, Inc.
- ACT, Inc. (1997c). *Setting achievement levels on the 1996 National Assessment of Educational Progress in Science: final report volume III, achievement levels-setting study*. Iowa City, IA: ACT, Inc.
- ACT, Inc. (1997d). *Setting achievement levels on the 1996 National Assessment of Educational Progress in Science: final report volume IV, validity evidence and special studies*. Iowa City, IA: ACT, Inc.
- ACT, Inc. (1997e). *Setting achievement levels on the 1996 National Assessment of Educational Progress in Science: final report volume V, technical decisions and NAGB actions*. Iowa City, IA: ACT, Inc.
- ACT, Inc. (1997f, March). *A proposal to the National Assessment Governing Board for developing achievement levels on the 1998 NAEP in Civics and Writing: Technical Proposal*. Iowa City, IA: ACT, Inc.
- ACT, Inc. (1997g, September). *Developing achievement levels on the 1998 NAEP in civics and writing: planning document*. Iowa City, IA: ACT, Inc.
- ACT, Inc. (1997h, December). *Developing achievement levels on the 1998 NAEP in civics and writing: design document*. Iowa City, IA: ACT, Inc.
- ACT, Inc. (1998, November). *Developing achievement levels on the 1998 NAEP in civics and writing: planning document, update 1*. Iowa City, IA: ACT, Inc.
- Bourque, M. L. (Ed.). (1995). *Proceedings of the joint conference on standard setting for large-scale assessments*. Washington, DC: National Assessment Governing Board, National Center for Education Statistics..

- Brennan, R. L. (1998). Misconceptions at the intersection of measurement theory and practice. *Educational Measurement: Issues and Practice*, 17(1), 5-9, 30.
- Burstein, L., Koretz, D. Linn, R., Sugrue, B., Novak, J., & Baker, E. L. (1996). Describing performance standards: Validity of the 1992 National Assessment of Educational Progress achievement level descriptions as characterizations of mathematics performance. *Educational Assessment*, 3(1), 9-51.
- Cizek, G. J. (1993). *Reaction to National Academy of Education Report, "Setting Performance Standards for Student Achievement."* Washington, DC: National Assessment Governing Board.
- DeMars, C. (1998). *The impact of test consequences and response format on performance.* Unpublished doctoral dissertation, Michigan State University, East Lansing, Michigan.
- Glaser, R., Linn, R., & Bohrnstedt, G. (1993). *Setting performance standards for student achievement.* Stanford, CA: National Academy of Education.
- Hambleton, R. K., & Bourque, M. L. (1991). *The levels of mathematics achievement (Volume III, Technical Report).* Washington, DC: National Assessment Governing Board.
- Joint Committee on Standards for Educational Evaluation. (1994). *The program evaluation standards: How to assess evaluations of educational programs (2nd ed).* Thousand Oaks, CA: Sage Publications.
- Kane, M. (1993). *Comment on the NAE evaluation of the NAGB achievement levels.* Washington, DC: National Assessment Governing Board.
- Kane, M. (1994). Validating the performance standards associated with passing scores. *Review of Educational Research*, 64(3), 425-461.
- Kane, M. (1995). Examinee-centered vs. task-centered standard setting. In M. L. Bourque (Ed.), *Joint Conference on Standard Setting for Large-Scale Assessments.* Washington: NCES-NAGB.
- Linn, R. L. (1999, February). *Assessment and accountability systems.* Invited presentation at the American Association of School Administrators, New Orleans.
- McLaughlin, D. H. (1993a). Validity of the 1992 NAEP achievement-level setting process. In R. Glaser, R. Linn, & G. Bohrnstedt (Eds.), *Setting performance standards for students achievement: Background studies.* Stanford, CA: National Academy of Education.
- McLaughlin, D. H. (1993b). Order of Angoff ratings in multiple simultaneous standards. In R. Glaser, R. Linn, & G. Bohrnstedt (Eds.), *Setting performance standards for students achievement: Background studies.* Stanford, CA: National Academy of Education.

- McLaughlin, D. H., DuBois, P. A., Eaton, M. S., Ehrlich, D. E., Stancavage, F. B., O'Donnell, C. A., Yu, J. Y., DeStefano, L., Pearson, D., Bottomley, D., Bullock, C. A., Hanson, M., & Rucinski, C. (1993). Comparison of teachers' and researchers' ratings for students' performance in mathematics and reading with NAEP measurement of achievement levels. In R. Glaser, R. Linn, & G. Bohrnstedt (Eds.), *Setting performance standards for students achievement: Background studies*. Stanford, CA: National Academy of Education.
- Mehrens, W. A. (1995). Methodological issues in standard setting for educational exams. In M. L. Bourque (Ed.), *Proceedings of the joint conference on standard setting for large-scale assessments*. Washington, DC: National Assessment Governing Board, National Center for Education Statistics.
- National Academy of Education. (1993). *Setting performance standards for student achievement*. Stanford, CA: National Academy of Education.
- Pellegrino, J. W., Jones, L. R., & Mitchell, K. J. (1998). *Grading the Nation's Report Card*. Washington, DC: National Academy Press.
- Reckase, M. D., & Bay, L. (1998, April). *Analysis of methods for collecting test-based judgements*. Paper presented at the meeting of the National Council on Measurement in Education, San Diego.
- Zieky, M. (1995). A historical perspective on setting standards. In M. L. Bourque (Ed.), *Joint Conference on Standard Setting for Large-Scale Assessments*. Washington: NCES-NAGB.

APPENDIX A

BIBLIOGRAPHY OF PUBLICATIONS, REPORTS, AND CONFERENCE PRESENTATIONS ON NAEP STANDARD-SETTING

Reference List

Publications

- Aldridge, P.W. (1994). The impact of standards on how teachers behave in the classroom: promises and perils. In M.L. Bourque (Ed.), *Proceedings of Joint Conference on Standard Setting for Large-Scale Assessments, Vol. II*. Washington, DC: National Assessment Governing Board.
- Berk, R.A. (1994). Standard setting—the next generation. In M.L. Bourque (Ed.), *Proceedings of Joint Conference on Standard Setting for Large-Scale Assessments, Vol. II*. Washington, DC: National Assessment Governing Board.
- Billeaud, K., Swygert, K., Nelson, L., & Thissen, D. (1998). Some ideas about item response theory applied to combinations of multiple-choice and open-ended Items: Scale scores for patterns of summed scores. In M.L. Bourque (Ed.), *Proceedings of Achievement Levels Workshop*. Washington, DC: National Assessment Governing Board.
- Bond, L. (1994). Ensuring fairness in the setting for performance standards. In M.L. Bourque (Ed.), *Proceedings of Joint Conference on Standard Setting for Large-Scale Assessments, Vol. II*. Washington, DC: National Assessment Governing Board.
- Bourque, M.L. (1999). The role of the National Assessment of Educational Progress (NAEP) in setting, reflecting, and linking national education policy to states' needs. In G.J. Cizek (Ed.), *Handbook of Educational Policy*, San Diego, CA: Academic Press.
- Bourque, M.L. (Ed.) (1998). *Proceedings of the National Assessment Governing Board Achievement Levels Committee Workshop*. Washington, DC: National Assessment Governing Board.
- Bourque, M.L. (Ed.) (1995). *Proceedings of the Joint Conference on Standard Setting for Large-Scale Assessments*. Washington, DC: National Assessment Governing Board and the National Center for Education Statistics.
- Bourque, M.L. (1993). The NAEP achievement level setting process for the 1992 mathematics assessment. In E.G. Johnson, J. Mazzeo, & D. Kline (Eds.), *Technical Report of the NAEP 1992 Trial State Assessment Program in Mathematics*. Washington, DC: National Center for Education Statistics.
- Bourque, M.L. (1993). The NAEP achievement level setting process for the 1992 reading assessment. In *Technical Report of the NAEP 1992 Trial State Assessment Program in Reading*. Washington, DC: National Center for Education Statistics.
- Bourque, M.L., Brinkley, M., & Atash, N. (1994). National standard setting and reporting. In T. Husen & T.N. Postlethwaite (Eds.), *The International Encyclopedia of Education, 2nd edition*. Oxford, England: Pergamon Press.

- Bourque, M.L., Champagne, A.B., & Crissman, S. (1997). *1996 Science Performance Standards: Achievement Results for the Nation and the States*. Washington, DC: National Assessment Governing Board.
- Bourque, M.L., & Garrison, H. (1991). *The STATE of Mathematics Achievement, Vols. I and II*. Washington, DC: National Assessment Governing Board
- Bourque, M.L., & Hambleton, R.K. (1993). Setting performance standards on the National Assessment of Educational Progress. *Measurement and Evaluation in Counseling and Development*, 26, 41-47.
- Brennan, R.L. (1998). Standard setting from the perspective of generalizability theory. In M.L. Bourque (Ed.), *Proceedings of Joint Conference on Standard Setting for Large-Scale Assessments, Vol. II*. Washington, DC: National Assessment Governing Board.
- Brennan, R.L. (1998). Misconceptions at the intersection of measurement theory and practice. *Educational Measurement: Issues and Practice*, 17 (1), 5-9, 30.
- Burstein, L., Koretz, D., Linn, R., Sugrue, B., Novak, J., Baker, E.L. & Lewis, E. (1996). Describing performance standards: Validity of the 1992 National Assessment of Educational Progress achievement level descriptors as characterizations of mathematics performance. *Educational Assessment* 3(1), 9-51.
- Collins, B.L. (1994). The consensus process in standards development. In Bourque, M.L. (Ed.), *Proceedings of Joint Conference on Standard Setting for Large-Scale Assessments, Vol. II*. Washington, DC: National Assessment Governing Board.
- Crocker, L. & Zieky, M. (1995). Executive summary of the Joint Conference on Standard Setting for Large-Scale Assessments. In M.L. Bourque (Ed.), *Proceedings of Joint Conference on Standard-Setting for Large-Scale Assessments*. Washington, DC: National Assessment Governing Board and the National Center for Education Statistics.
- Forsyth, R.A. (1998). NAEP frameworks and achievement levels. In Bourque, M.L. (Ed.), *Proceedings of Achievement Levels Workshop*. Washington, DC: National Assessment Governing Board.
- Goldberg, J.P. (1998). The recommended dietary allowances: Can they inform the development of standards of academic achievement? *Applied Measurement of Education*, 11 (1), 97-105.
- Hambleton, R.K. & Bourque, M.L. (1991). *The LEVELS of Mathematics Achievement, Vol. III*. Washington, DC: National Assessment Governing Board.

- Hambleton, R.K. & Slater, S. (1994). Using performance standards to report national and state assessment data: are the reports understandable and how can they be improved? In M.L. Bourque (Ed.), *Proceedings of Joint Conference on Standard Setting for Large-Scale Assessments, Vol. II*. Washington, DC: National Assessment Governing Board.
- Hänsler, L.M. (1998). Lessons for the National Assessment of Educational Progress from military standard setting. *Applied Measurement in Education*, 11 (1), 81-95.
- Huynh, H. (1994). Some technical aspects of standard setting. In Bourque, M.L. (Ed.), *Proceedings of Joint Conference on Standard Setting for Large-Scale Assessments, Vol. II*. Washington, DC: National Assessment Governing Board.
- Jaeger, R.M. (1994). On the cognitive construction of standard-setting judgements: the case of configural scoring. In M.L. Bourque (Ed.), *Proceedings of Joint Conference on Standard Setting for Large-Scale Assessments, Vol. II*. Washington, DC: National Assessment Governing Board.
- Jaeger, R., Mullis, I., Bourque, M.L., & Shakrani, S. (1996). Setting performance standards for performance assessments: some fundamental issues, current practice, and technical dilemmas. In G. Phillips (Ed.), *Technical Issues in Performance Assessment*. Washington, DC: National Center for Education Statistics.
- Jasanoff, S. (1998). Science and judgement in environmental standard setting. *Applied Measurement in Education*, 11 (1), 107-120.
- Kane, M. (1994a). Examinee-centered vs. task-centered standard setting. In M.L. Bourque (Ed.), *Proceedings of Joint Conference on Standard Setting for Large-Scale Assessments, Vol. II*. Washington, DC: National Assessment Governing Board.
- Kane, M. (1994b). Validating the performance standards associated with passing scores. *Review of Educational Research* 64, 425-461.
- Koretz, D.M., & Deibert, E. (1995/1996). Setting standards and interpreting achievement: A cautionary tale from the National Assessment of Educational Progress. *Educational Assessment* 3(1), 53-81.
- Linn, R.L. (1994). The likely impact of performance standards as a function of uses: from rhetoric to sanctions. In M.L. Bourque (Ed.), *Proceedings of Joint Conference on Standard Setting for Large-Scale Assessments, Vol. II*. Washington, DC: National Assessment Governing Board.
- Linn, R.L. (1998). Validating inferences from National Assessment of Educational Progress achievement level reporting. *Applied Measurement in Education* 11 (1), 23-47.
- Lissitz, R. & Bourque, M.L. (1995). Reporting NAEP results using standards. *Educational Measurement: Issues and Practice*, 14, 14-23.

- Livingston, S.A. (1994). Standards for reporting the educational achievement of groups. In M.L. Bourque (Ed.), *Proceedings of Joint Conference on Standard Setting for Large-Scale Assessments, Vol. II*. Washington, DC: National Assessment Governing Board.
- McLaughlin, D.H., DuBois, P.A., Eaton, M.S., Ehrlich, D.E., Stancavage, F.B., O'Sonnoell, C.A., Yu, J.Y., DeStefano, L., Pearson, D. Bottomley, D., Bullock, C.A., Hanson, M. & Rucinski, C. (1993). Comparison of teachers' and researchers' rating of students' performance in mathematics and reading with NAEP measurement of achievement levels. In R. Glazer, R. Linn., & G. Bohrnstedt (Eds.), *Setting Performance Standards for Students Achievement: Background Studies*. Stanford, CA: National Academy of Education.
- Mehrens, W.A. (1994). Methodological issues in standard setting for educational exams. In M.L. Bourque (Ed.), *Proceedings of Joint Conference on Standard Setting for Large-Scale Assessment, Vol. II*. Washington, DC: National Assessment Governing Board.
- Messick, S. (1994). Standards-based score interpretation: establishing valid grounds for valid inferences. In M.L. Bourque (Ed.), *Proceedings of Joint Conference on Standard Setting for Large-Scale Assessments, Vol. II*. Washington, DC: National Assessment Governing Board.
- Mislevy, R. (1998). Implications of market-basket reporting for achievement level setting. *Applied Measurement in Education*, 11 (1), 49-63.
- Moss, P.A. (1994). An interpretive approach to setting and applying standards. In M.L. Bourque (Ed.), *Proceedings of Joint Conference on Standard Setting for Large-Scale Assessments, Vol. II*. Washington, DC: National Assessment Governing Board.
- National Academy of Education (1992). *Assessing student achievement in the states*. Stanford, CA: Author.
- National Academy of Education (1993a). *Setting performance standards for student achievement*. Stanford, CA: Author.
- National Academy of Education (1993b). *The trial state assessment: Prospects and realities*. Stanford, CA: Author.
- National Academy of Education (1996). *Quality and utility: The 1994 trial state assessment in reading*. Stanford, CA: Author.
- National Academy of Education. (1997). *Assessment in transition: Monitoring the nation's educational progress*. Stanford, CA: Author.
- National Academy of Sciences (1998). *Grading the nation's report card: Evaluating NAEP and transforming the Assessment of Educational Progress*. Washington, DC: Author.

- National Assessment Governing Board (1997). *What Do Students Know? 1996 NAEP Science Results for 4th, 8th, and 12th, Graders*. Washington, DC: Author.
- Pellegrino, J., Wise, L., & Raju, N. (1998). Guest editors' note. *Applied Measurement in Education*, 11 (1), 1-7.
- Phillips, S.E. (1994). Legal defensibility of standards: issues & policy perspectives. In M.L. Bourque (Ed.), *Proceedings of Joint Conference on Standard Setting for Large-Scale Assessments, Vol. II*. Washington, DC: National Assessment Governing Board.
- Phillips, G.W., Mullis, I.V.S., Bourque, M.L., Williams, P.L., Hambleton, R.K., Owens, E.H., & Barton, P.E. (1993). *Interpreting NAEP Scales*. Washington, DC: National Center for Education Statistics.
- Plake, B. (1998). Setting performance standards for professional licensure and certification. *Applied Measurement in Education*, 11 (1), 65-80.
- Popham, W.J. (1998). The criticality of consequences in standard setting: six lessons learned the hard way by a standard-setting abettor. In M.L. Bourque (Ed.), *Proceedings of Achievement Levels Workshop*. Washington, DC: National Assessment Governing Board.
- Reckase, M.D. (1998). Converting boundaries between National Assessment Governing Board performance categories to points on the National Assessment of Educational Progress score scale: The 1996 science NAEP process. *Applied Measurement in Education*, 11 (1), 9-21.
- Shepard, L.A. (1994). Implications for standard setting of the National Academy of Education evaluation of the National Assessment of Educational Progress achievement levels. In M.L. Bourque (Ed.), *Proceedings of Joint Conference on Standard Setting for Large-Scale Assessments, Vol. II*, Washington, DC: National Assessment Governing Board.
- Smith, H.W. (1994). Issues in standards setting: making a decentralized system responsive to centralized policy making. In M.L. Bourque (Ed.), *Proceedings of Joint Conference on Standard Setting for Large-Scale Assessments, Vol. II*. Washington, DC: National Assessment Governing Board.
- Thissen, D., Billeaud, K., McLeod, L., & Nelson, L. (1998). A brief introduction to item response theory for items scored in more than two categories. In M.L. Bourque (Ed.), *Proceedings of Achievement Levels Workshop*. Washington, DC: National Assessment Governing Board.
- van der Linden, W.J. (1994). A conceptual analysis of standard setting in large-scale assessments. In M.L. Bourque (Ed.), *Proceedings of Joint Conference on Standard Setting for Large-Scale Assessments, Vol. II*. Washington, DC: National Assessment Governing Board.
- van der Linden, W.J. (1998). Assembly of test forms for use in large-scale educational assessment. In M.L. Bourque (Ed.), *Proceedings of Achievement Levels Workshop*.

Washington, DC: National Assessment Governing Board.

Zieky, M.J. (1994). A historical prospective on setting standards. In M.L. Bourque (Ed.), *Proceedings of Joint Conference on Standard Setting for Large-Scale Assessments, Vol. II*. Washington, DC: National Assessment Governing Board.

Technical Reports and Correspondence

- Bay, L. (1997, July). *An investigation of context effects based on item-block position on the 1996 NAEP science for grade 8*. Unpublished manuscript.
- Bay, L. (1998, February). *1998 NAEP Writing Achievement Levels-Setting Process performance profiles study*. A paper presented at the meeting of the Technical Advisory Committee for Standard Setting (TACSS), St. Louis, MO.
- Bay, L., Chen, W.H., & Loomis, S.C. (1997, May). *Observations about the 1996 NAEP Science Assessment*. Unpublished manuscript.
- Bay, L., Chen, L., & Reckase, M.D. (1997, October). *The grid: A possible rating method for the 1998 NAEP writing achievement levels-setting process*. A paper presented at the meeting of the Technical Advisory Committee for Standard Setting (TACSS), St. Louis, MO.
- Bay, L., & Hanson, B.A. (1997, October). *Computing achievement levels cutpoints from NAEP booklet classification studies: A secondary analysis*. A paper presented at the meeting of the Technical Advisory Committee for Standard Setting (TACSS), St. Louis, MO.
- Bay, L., & Nicewander, W.A. (1997). *Reliability estimates of polytomous items*. Unpublished manuscript.
- Beaton, A.E. (Nov 7, 1989). *Setting NAEP performance goals*. Personal communication.
- Bock, R. D. (1997). *Contextual validity of the NAGB science achievement levels (draft)*. Unpublished manuscript.
- Brennan, R.L. (1993). Writing reliabilities and disattenuated correlations for grade 12. In *Setting achievement levels on the 1992 National Assessment of Educational Progress in mathematics, reading, and writing: A Technical report on reliability and validity*. Iowa City, IA: ACT.
- Burstein, L., Koretz, D.M., Linn, R.L., Sugrue, B., Novak, J., Lewis, E., & Baker, E.L. (1993). *The validity of interpretations of the 1992 NAEP achievement levels in mathematics*. Los Angeles, CA: Center for the Study of Evaluation, University of California
- Cohen, J., Greenberg, E., & Bohrnstedt, G. (1997). *NAEP achievement standards: Background and options*. Washington, DC: American Institutes for Research.
- Cizek, G.J. (1993). Reactions to National Academy of Education report, *Setting Performance Standards for Student Achievement*. (Available from the National Assessment Governing Board, Washington, DC).

- Hambleton, R.K. (Dec 13, 1990). *Recommended adjustments in the grades 4, 8, and 12 achievement levels*. Personal communication.
- Hambleton, R.K. (1993). *Setting achievement levels on the 1990 NAEP mathematics assessment: Response to technical criticisms*. (Laboratory of Psychometric and Evaluative Research Report No. 250). Amherst, MA: University of Massachusetts.
- Hambleton, R.K., & Bourque, M.L. (1993). *Setting achievement levels on the 1990 mathematics assessment* (Laboratory of Psychometric and Evaluative Research Report No. 217). Amherst, MA: University of Massachusetts,
- Kane, M. (1993a). *Comment on the NAE evaluation of the NAGB achievement levels*. (Available from the National Assessment Governing Board, Washington, DC).
- Kane, M. (1993b). *The validity of performance standards*. (Available from the National Assessment Governing Board, Washington, DC).
- Linn, R.L., Koretz, D.M., Baker, E.L., & Burstein, L. (1991). *The validity and credibility of the achievement levels for the 1990 National Assessment of Educational Progress in mathematics*. Los Angeles, CA: Center for the Study of Evaluation, University of California.
- Linn, R.L., Koretz, D.M., Baker, E.L., & Burstein, L. (1996). *Assessing the validity of the National Assessment of Educational Progress: NAEP Technical Review Panel White Paper*. Los Angeles, CA: Center for the Study of Evaluation, University of California
- Luecht, R.M. (1992). Interpolating Θ for the partial credit model. In *Setting Achievement Levels on the 1992 National Assessment of Educational Progress in Mathematics, Reading and Writing: A Technical Report on Reliability and Validity*. Iowa City, IA: ACT.
- Luecht, R.M. (1993a). Some results from the NAGB ALS project in reading. In *Setting Achievement Levels on the 1992 National Assessment of Educational Progress in Mathematics, Reading, and Writing: A Technical report on Reliability and Validity*. Iowa City, IA: ACT.
- Luecht, R.M. (1993b). NAGB ALS Project in writing: preliminary mapping results. In *Setting achievement Levels on the 1992 National Assessment of Educational Progress in Mathematics, Reading, and Writing: A Technical report on Reliability and Validity*. Iowa City, IA: ACT.
- Luecht, R.M. (1993c). Using IRT to improve the standard setting process for dichotomous and polytomous items. In *Setting Achievement Levels on the 1992 National Assessment of Educational Progress in Mathematics, Reading, and Writing: A Technical report on Reliability and Validity*. Iowa City, IA: ACT.
- Luecht, R.M. (1993d). Some results on the stability of the NAEP achievement level standards in

mathematics. In *Standard Setting Achievement Levels on the 1992 National Assessment of Educational Progress in Mathematics, Reading, and Writing: A Technical report on reliability and Validity*. Iowa City, IA: ACT.

National Assessment Governing Board (1991). *Response to the draft summative evaluation report on the National Assessment Governing Board's inaugural effort to set achievement levels on the National Assessment of Educational Progress*. Washington, DC: Author.

National Assessment Governing Board (1992). *Survey of reactions to the use of achievement levels in reporting 1990 NAEP Mathematics Results*. Unpublished manuscript.

Stufflebeam, D.L., Jaeger, R.M., & Scriven, M. (1991, August). *Summative evaluation of the National Assessment Governing Board's inaugural 1990-91 effort to set achievement levels on the National Assessment of Educational Progress*. (Available from the National Assessment Governing Board, Washington, DC).

U.S. General Accounting Office (1993). Educational achievement standards: NAGB's approach yields misleading interpretations. (GAO/PEMD-93-12). Washington, DC: Author.

ACT Technical Reports on NAEP Achievement Level Setting, 1991 -1999

- ACT. (1991). *Planning document for setting achievement levels on the 1992 National Assessment of Educational Progress in mathematics, reading, and writing*. Iowa City, IA: Author.
- ACT. (1992a). *Design document for setting achievement levels on the 1992 National Assessment of Educational Progress in mathematics, reading, and writing*. Iowa City, IA: Author.
- ACT. (1992b). *Setting achievement levels on the 1992 National Assessment of Educational Progress: pilot study report*. Iowa City, IA: Author.
- ACT. (1993a). *Setting achievement levels on the 1992 National Assessment of Educational Progress in reading: Final report*. Iowa City, IA: Author.
- ACT. (1993b). *Description of mathematics achievement levels process and proposed achievement levels descriptions. National Assessment of Educational Progress in mathematics: Final report*. Iowa City, IA: Author.
- ACT. (1993c). *Setting achievement levels on the 1992 National Assessment of Educational Progress in writing: Final report*. Iowa City, IA: Author.
- ACT. (1993d). *Setting achievement levels on the 1994 National Assessment of Educational Progress in U.S. history, geography, and science: Design document*. Iowa City, IA: Author.
- ACT. (1993e). *Setting achievement levels on the 1992 National Assessment of Educational Progress in reading, mathematics, and writing: A technical report on reliability and validity*. Iowa City, IA: Author.
- ACT. (1994a). *Planning document for developing achievement levels for the 1994 National Assessment of Educational Progress in the U.S. history and world geography assessments (and future science assessment)*. Iowa City, IA: Author.
- ACT. (1994b). *Setting achievement levels on the 1994 National Assessment of Educational Progress in geography and U.S. history and the 1996 National Assessment of Educational Progress in science: Design document*. Iowa City, IA: Author.
- ACT. (1994c). *Results of the 1994 U.S. history National Assessment of Educational Progress achievement levels-setting pilot study*. Iowa City, IA: Author.
- ACT. (1994d). *Results of the 1994 geography National Assessment of Educational Progress achievement levels-setting pilot study*. Iowa City, IA: Author.

- ACT. (1995a). *Preliminary report on the 1994 National Assessment of Educational Progress achievement levels-setting process for U.S. history, and geography*. Iowa City, IA: Author.
- ACT. (1995b). *NAEP reading revisit: An evaluation of the 1992 achievement levels descriptions*. Iowa City, IA: Author.
- ACT. (1997a). *Preliminary report on the 1996 achievement levels-setting process for science*. Iowa City, IA: Author.
- ACT. (1997b). *Developing achievement levels on the 1998 NAEP in civics and writing: Design document*. Iowa City, IA: Author.
- ACT. (1997c). *Developing achievement levels on the 1998 NAEP in civics and writing: Planning document*. Iowa City, IA: Author.
- ACT. (1997d). *Setting achievement levels on the 1996 National Assessment of Educational Progress in science. Final report, Executive summary*. Iowa City, IA: Author.
- ACT. (1997e). *Setting achievement levels on the 1996 National Assessment of Educational Progress in science. Final report, Volume I. Pilot study 1*. Iowa City, IA: Author.
- ACT. (1997f). *Setting achievement levels on the 1996 National Assessment of Educational Progress in science. Final report, Volume II. Pilot study 2*. Iowa City, IA: Author.
- ACT. (1997g). *Setting achievement levels on the 1996 National Assessment of Educational Progress in science. Final report, Volume III. Achievement levels-setting study*. Iowa City, IA: Author.
- ACT. (1997h). *Setting achievement levels on the 1996 National Assessment of Educational Progress in science: Final report, Volume IV. Validity evidence special studies*. Iowa City, IA: Author.
- ACT. (1997i). *Setting achievement levels on the 1996 National Assessment of Educational Progress in science. Final report, Volume V. Technical decisions NAGB actions*. Iowa City, IA: Author.
- ACT. (1999a). *Developing achievement levels on the 1998 NAEP in civics: Field trial report*. Iowa City, IA: Author.
- ACT. (1999b). *Developing achievement levels on the 1998 NAEP in writing: Field trial report*. Iowa City, IA: Author.
- ACT. (1999c). *Developing achievement levels on the 1998 NAEP civics: Pilot study report*. Iowa City, IA: Author.

ACT. (1999d). *Developing achievement levels on the 1998 NAEP writing: Pilot study report*.
Iowa City, IA: Author.

Conference Presentations

- Bay, L. (1996, April). *Psychometric issues in setting standards on the NAEP*. Paper presented at the meeting of the National Council on Measurement in Education, New York, NY
- Bay, L. (1998, April). *Comparing student performance on different item formats relative to achievement levels cutpoints*. Paper presented at the meeting of the National council on Measurement in Education, San Diego, CA.
- Bay, L., & Loomis, S.C. (1995, June). *Validation of the 1994 NAEP achievement levels: Booklet classification study*. Paper presented at the Large-Scale Assessment Conference, Phoenix, AZ.
- Bay, L. & Nering, M. (1998, April). *A demonstration of using person-fit statistics in standard setting*. Paper presented at the meeting of the National Council for Measurement in Education, San Diego, CA.
- Bourque, M.L. (1993a, June). *Policy issues in setting achievement standards on assessment*. Paper presented at the CCSSO National Conference on Large Scale Assessment, Albuquerque, NM.
- Bourque, M.L. (1993b, April). *The NAEP achievement level setting process for the 1992 mathematics assessment*. Paper presented at the meeting of the National Council on Measurement in Education, Atlanta, GA.
- Bourque, M.L. (1993c, April). *The NAEP achievement level setting process for the 1992 reading assessment*. Paper presented at the 38th Annual Convention of the International Reading Association Convention, San Antonio, TX.
- Bourque, M.L. (1993d, June). *Update on NAEP achievement levels*. Paper presented at the CCSSO National Conference on Large Scale Assessment, Albuquerque, NM.
- Bourque, M.L. (1994, June). *Setting standards on the 1994 NAEP*. Paper presented at the CCSSO National Conference on Large Scale Assessment, Albuquerque, NM.
- Bourque, M.L. (1995, August). *Setting student performance standards on the National Assessment of Education Progress*. Paper presented at the meeting of the American Psychological Association, New York City, NY.
- Bourque, M.L., & Loomis, S.C. (1996, April). *Psychometric considerations of constructed response items for setting achievement levels on the NAEP*. Paper presented at the meeting of the National Council on Measurement in Education, New York City, NY.
- Bourque, M.L., & Truby R. (1991, June). *Setting achievement levels for reporting NAEP results*. Paper presented at the ECS/CDE Large-Scale Assessment Conference, Boulder, CO.

- Bourque, M.L., & Webb, II, M. W. (1992, June). *How have NAEP proficiency scales been improved?* Paper presented at the ECS/CDE Large-Scale Assessment Conference, Boulder, CO.
- Bourque, M.L., & Webb II, M.W. (1993, April). *The 1992 NAEP mathematics achievement levels.* Paper presented at the meeting of the National Council on Measurement in Education, Atlanta, GA.
- Burstein, L. (1994, April). *Achievement levels reporting on the National Assessment of Educational Progress: A technical and policy debate.* Paper presented at the meeting of the National Council on Measurement in Education, New Orleans, LA.
- Chen, W.H. (1996, April). *Psychometric issues in setting standards on the NAEP.* Paper presented at the meeting of the National Council on Measurement in Education, New York, NY.
- Chen, W.H. (1998a, April). *Methods for combining item-by-item ratings.* Paper presented at the meeting of the National Council on Measurement in Education, San Diego, CA.
- Chen, W.H. (1998b, April). *Setting achievement level standards for NAEP using response pattern estimation: a simulation study.* Paper presented at the meeting of the National Council on Measurement in Education, San Diego, CA.
- Davey, T. (1996, April). *Psychometric issues in setting standards on the NAEP.* Paper presented at the meeting of the National Council on Measurement in Education, New York, NY.
- Fan, M., & Hsu, Y. (1997, April). *Comparison of different judgmental standard-setting procedures for NAEP data.* Paper presented at the meeting of the American Educational Research Association, Chicago, IL.
- Hambleton, R.K., & Bourque, M.L. (1992, April). *Methodological considerations in setting standards on national examinations.* Invited paper presented at the meeting of the American Educational Research Association, San Francisco, CA.
- Hanson, B.A. & Bay, L.G. (1999, April). *Classifying student performance as a method for setting achievement levels for NAEP writing.* Paper presented at the meeting of the National Council on Measurement in Education, Montreal, Canada.
- Hanson, B.A., Bay, L., & Loomis, S.C. (1998, April). *Booklet classification study.* Paper presented at the meeting of the National Council on Measurement in Education, San Diego, CA.
- Hanick, P. L. & Loomis, S.C. (1999, April). *Setting standards for the 1998 NAEP in civics and writing: Using focus groups to finalize the achievement levels descriptions.* Paper

presented at the meeting of the National Council on Measurement in Education, Montreal, Canada.

Johnson, E. (1996, April). *Psychometric issues in setting standards on the NAEP*. Paper presented at the meeting of the National Council on Measurement in Education, New York, NY.

Johnson, E.G., Liang, J.L., Norris, N., Rogers, A., & Nicewander, A. (1996, April). *Directly estimated NAEP scale values from double-length assessment booklets--a replacement for plausible values*. Paper presented at the meeting of the National Council on Measurement in Education, New York City, NY.

Kane, M. (1996, April). *Psychometric issues in setting standards on the NAEP*. Paper presented at the meeting of the National Council on Measurement in Education, New York, NY.

Kane, M.T., & Bay, L. (1996, April). *The intersection of score interpretation, assessment design, and standard-setting methodology on NAEP*. Paper presented at the meeting of the National Council on Measurement in Education, New York.

Linn, R.L., & Shepard, L.A. (1997, June). *Item-by-item standard setting: Unintended consequences due to less than perfect item-correlations*. Paper presented at the Large Scale Assessment Conference, Colorado Springs, CO.

Loomis, S.C. (1998, April). *Setting achievement levels for the 1996 science NAEP*. Paper presented at the meeting of the National Council on Measurement in Education, San Diego, CA.

Loomis, S.C., & Bay, L. (1995, October). *Student performance on the 1994 NAEP: geography achievement levels*. Paper presented at the meeting of the National Council for Geographic Education, San Antonio, TX.

Loomis, S.C., Bourque, M.L., & Webb II, M.W. (1993, April). *Enhancements to the process of setting achievement levels for the 1992 NAEP in mathematics, reading, and writing*. Paper presented at the meeting of the National Council on Measurement in Education, Atlanta, GA.

Luecht, R.M. (1993a, April). *Using IRT to improve the standard setting process for dichotomous and polytomous items*. Paper presented at the meeting of the National Council on Measurement in Education, Atlanta, GA.

Luecht, R.M. (1993b, April). *Some results on the stability of the NAEP achievement level standards in mathematics*. Paper presented at the meeting of the National Council on Measurement in Education, Atlanta, GA.

McColl, G. (1994, April). *Achievement levels reporting on the National Assessment of Educational Progress: A technical and policy debate*. Paper presented at the meeting of

the National Council on Measurement in Education, New Orleans, LA.

Nicewander, A. (1996, April). *Psychometric issues in setting standards on the NAEP*. Paper presented at the meeting of the National Council on Measurement in Education, New York, NY.

Reckase, M.D. (1996a, April). *Psychometric issues in setting standards on the NAEP*. Paper presented at the meeting of the National Council on Measurement in Education, New York, NY.

Reckase, M.D. (1998, April). *Methods for collecting test-based judgements*. Paper presented at the meeting of the National Council on Measurement in Education, San Diego, CA.

Reckase, M.D., Loomis, S.C., & Wang, T. (1995, April). *Setting standards on NAEP performance items*. Paper presented at the meeting of the National Council on Measurement in Education, San Francisco, CA.

Reckase, M.D., Loomis, S.C., Wang, T., & Bay, L. (1995, April). *Setting standards on NAEP performance items*. Paper presented at the meetings of the American Educational Research Association and the National Council on Measurement in Education, San Francisco, CA.

Saterfiel, T. (1994, April). *Achievement levels reporting on the National Assessment of Educational Progress: A technical and policy debate*. Paper presented at the meeting of the National Council on Measurement in Education, New Orleans, LA.

Shepard, L. (1994, April). *Achievement levels reporting on the National Assessment of Educational Progress: A technical and policy debate*. Paper presented at the meeting of the National Council on Measurement in Education, New Orleans, LA.

Skaggs, G., & Lissitz, B.W. (1990, April). *Equating NAEP with state, local, commercial or international assessments*. Paper presented at the meeting of the Council of Chief State School Officers, Washington, DC.

Webb II, M.W., & Loomis, S. (1993, April). *Results of content validity studies related to the 1992 NAEP achievement levels for mathematics, reading, and writing*. Paper presented at the meeting of the National Council on Measurement in Education, Atlanta, GA.

The Honorable William Jefferson Clinton
President
The White House
1600 Pennsylvania Avenue
Washington, DC 20500

91-6-11-1-16

OFFICE OF GOVERNMENT & NATIONAL AFFAIRS
WASHINGTON, DC



Director
JESS N. HORDS

Assistant Director
STACY BURDETT

Washington Counsel
MICHAEL LIEBERMAN

NATIONAL COMMISSION
OFFICERS

National Chair
HOWARD P. BERKOWITZ

National Director
ABRAHAM H. FOXMAN

Chair, National
Executive Committee
GLEN A. TOBIAS

Associate National Director
Director, Development
PETER T. WILLNER

Chair, Washington Affairs
Committee
HAROLD GERSHOWITZ

December 23, 1999

The Honorable Bruce N. Reed
Assistant to the President for Domestic Policy
The White House
Washington, DC 20500

Dear Mr. ^{Bruce} Reed:

I wanted to make sure that you saw a copy of the attached.

With every best wish.

Sincerely,

A handwritten signature in black ink, appearing to read "Jess N. Hordes", written over a horizontal line.

Jess N. Hordes
Washington Representative

JNH:mes
Encl.

Andy -
FYI
BR



OFFICERS OF THE
NATIONAL COMMISSION

National Chair
HOWARD P. BERKOWITZ

National Director
ABRAHAM H. FOXMAN

Chair
National Executive Committee
GLEN A. TOBIAS

Associate National Director
Director, Development
PETER T. WILLNER

Honorary Chairs
KENNETH J. BIALKIN
SEYMOUR GRAUBARD
MAXWELL E. GREENBERG
BURTON M. JOSEPH
BURTON S. LEVINSON
MELVIN SALBERG
DAVID H. STRASSLER

Vice Chairs
MEYER EISENBERG
THOMAS C. HOMBERGER
JUDITH KRUPP
LESTER POLLACK
ALVIN J. ROCKOFF
ROBERT G. SUGARMAN

Honorary Vice Chairs
LEONARD L. ABESS
DOROTHY BINSTOCK
RUDY BOSCHWITZ
EDGAR M. BRONFMAN
MAXWELL DANI
MAX FISHER
BRUCE L. HOCHMAN
SYDNEY P. JAKKOW
GERI M. JOSEPH
MAX M. KAMPELMAN
SAM KANE
PHILIP M. KLUTZNICK
PHILIP KRUPP
BERNARD MARCUS
SAMUEL H. MILLER
BERNARD D. MINTZ
MILTON MOLLEN
ROBERT R. NATHAN
ROWLAND SCHALFER
IRVING SHAPIRO
JOEL SPRAYREGEN
WILLIAM VIPRIN
EUGENE WARNER
CARMEN WARSCHAW
SIDNEY R. YATES

Vice Chair
National Executive Committee
GEORGE L. MOSS

Honorary Chair
National Executive Committee
RONALD B. SOBEL

Treasurer
ROBERT H. NAFTALY

Assistant Treasurer
MURRAY KOPPELMAN

Secretary
BARBARA B. BAUSER

Assistant Secretary
I. BARRY MEHLER

President, B'nai B'rith
RICHARD D. HEIDEMAN

Executive Vice President
B'nai B'rith
SIDNEY CLEARFIELD

Assistant National Director
Director, International Affairs
KENNETH JACOBSON

DIVISION DIRECTORS
Civil Rights
ELIZABETH I. COLEMAN

Community Service
ANN TOURK

Education
CARYL STERN-LAROSA

Finance & Administration
BOBBIE ARBESFELD

Leadership
Assistant to the National Director
MARK D. MEDIN

Marketing & Communications
MARK A. EDELMAN

Washington Representative
JESS N. HORDS

General Counsel
ARNOLD FORSTER

December 22, 1999

The President
The White House
Washington, DC 20500

Dear Mr. President:

We commend your thoughtful radio address last Saturday, and support your administration's continuing efforts to provide guidance for school officials on permissible religious expression in public schools. At the same time, however, we have some concerns about the Department of Education initiative to encourage expanded partnerships between public schools and religious institutions.

As you know, the Anti-Defamation League has long believed that a strong separation of church and state is essential to preserving and promoting religious rights and liberties in our increasingly pluralistic nation. Like you, however, we have never believed that the Constitution requires public schools to be "religion-free zones." To this end, we were pleased to serve on the Drafting Committee for the 1995 publication Religion in the Public Schools: A Joint Statement of Current Law, the basis for Secretary Riley's original guidelines for public schools.

School officials have relied on these guidelines, and we have no doubt that the new Department of Education mailing you discussed in your radio address will provide them with practical guidance to safeguard students' private religious expression while avoiding religious coercion and any appearance of state sponsorship of religious activities.

Unfortunately, we do not share the same enthusiasm for the initiative you announced to promote expanded partnerships between public schools and religious institutions. School-church partnerships lack the consensus support across the ideological spectrum of your earlier guidelines, and their legality is far from settled.

When we discussed this concern with representatives of the Department of Education, they agreed that the Department should prepare a cover letter to principals offering some clarifications. Secretary Riley's "Guidelines for School Officials, Volunteers and Mentors Participating in Public School Community Partnerships," a set of practical "do's" and "don'ts" intended to help establish appropriate boundaries, uphold constitutional protections, and avoid coercive arrangements, was an additional welcome response to our concerns, and school officials will no doubt find them useful. Nevertheless, we remain troubled by the potential for abuse in these school-church partnerships.

The President
December 22, 1999
Page 2

Mr. President, we share your view that our nation's "remarkable religious diversity" is one of America's greatest sources of strength. We have greatly admired your administration's leadership efforts on behalf of religious freedom in this country -- and across the globe. We appreciate your consideration of the concerns we have expressed in this letter, and look forward to continuing to work with you to safeguard our precious religious freedom.

Sincerely,


Howard P. Berkowitz
National Chairman


Abraham H. Foxman
National Director



Cynthia A. Rice

03/14/2000 10:05:08 AM

Record Type: Record

To: William Marshall/WHO/EOP, Edward W. Correia/WHO/EOP

cc: Andrea Kane/OPD/EOP, Eugenia Chough/OPD/EOP

Subject: Comments on C Choice Guidelines

We have two comments on the charitable choice guidelines, related to the programmatic application of these principles:

1. Page 1, first paragraph, 2nd sentence: reference to "direct federal aid." It's unclear to us if this includes federal funds which flow through states, localities, or other entities (such as nonprofits) as is common in many of our major programs (e.g., the TANF welfare block grant, Welfare-to-Work formula grants, the SAMHSA substance abuse block grant, the child care block grant, etc.), but we assume it does. Perhaps you could clarify this by adding "..such "faith based" organizations are eligible for direct federal aid, including direct federal grants and federal funds provided through state or local governments or other entities . . ."

2. Page 3, first paragraph, sentence beginning "For example." This sentence strikes us as a fairly detailed definition of "religious worship, religious instruction, or proselytizing," and since the footnote doesn't seem to provide a relevant case citation, I take it this is a definition or example devised by OLC. Given our recent discussions about the value of leaving constitutional interpretation to the courts, this level of detail makes us a bit uncomfortable. How about replacing that sentence with one that reads something like "This is true even in cases where religious worship, religious instruction, or proselytizing helps achieve a secular goal, such as discouraging teen pregnancy, crime, or substance abuse."

Draft 2/25/00

Internal

**Charitable Choice
General Guidelines for Evaluating Policy Proposals**

Government programs that provide support, such as grants, to private organizations for the purpose of assisting or encouraging such organizations to help solve social problems may provide direct aid to religious organizations that are not pervasively sectarian. Consequently, the Administration has made efforts to ensure that, in an array of social service programs, such “faith-based” organizations are eligible for direct federal aid, including federal grants, on the same terms and conditions as other community and nonprofit organizations. However, there are some constitutional limitations that must be kept in mind.

1. Even where the recipient is using governmental aid in a substantially secular setting, the government must ensure that the aid is not used to advance religious activities — such as religious worship, religious instruction, proselytizing, or the financing of the construction or maintenance of facilities in which such religious activities would take place.
2. The government may not provide financial aid directly to “pervasively sectarian” institutions, defined as institutions in which religion is so pervasive that a substantial portion of their functions are subsumed in the religious mission. Churches and church elementary and secondary schools are examples of organizations that traditionally have been deemed pervasively sectarian. Other religious organizations must be evaluated on a case-by-case basis to determine whether they are pervasively sectarian.
3. The criteria for funding should be neutral and secular. The government may not provide aid to a particular organization because it is, or is not, religious. Accordingly, government aid should not be limited to religious organizations in general, and the government must not otherwise prefer such organizations over others, e.g., by setting aside a particular portion of funds for them. Similarly, a secular organization should not be favored for aid over a nonpervasively sectarian religious organization because of the latter’s religious character or affiliations.
4. The government may not prefer certain religious organizations over others, except on the basis of secular criteria unrelated to the organizations’ religious affiliations or beliefs.

Explanatory Notes to Charitable Choice Guidelines

The President and the Vice President each have stated that faith-based organizations can play an important role in providing social services. And the Supreme Court has held that, pursuant to a governmental program that provides governmental support, such as grants, to private organizations for the purpose of assisting or encouraging such organizations to help solve social problems, the government may provide direct aid to religious organizations that are not pervasively sectarian. See Bowen v. Kendrick, 487 U.S. 589 (1988). Consequently, the Administration has supported efforts to ensure that, in an array of social service programs, such organizations are eligible for federal aid, including federal grants, on the same terms and conditions as other community and nonprofit organizations.¹

Kendrick and other cases establish that an institution's religious affiliations do not constitutionally disqualify it from participating equally in a neutral, governmental financial-aid program that benefits both religious and nonreligious entities.² However, there are some constitutional limitations that must be kept in mind. The following guidelines provide a general framework of the most important Establishment Clause restrictions. Nevertheless, because the constitutional doctrine in this area is somewhat complex and ambiguous, and because the constitutionality of a particular form of government aid to religious organizations might depend significantly on the particular statutory or regulatory scheme and on the characteristics of recipient organizations, agencies should consult with their general counsels and, where appropriate, with the Department of Justice Office of Legal Counsel, before embarking on a program of aid to such organizations.

1. The government must ensure that direct government aid is not used to advance “

¹ The Department of Justice previously has noted that bills introduced in Congress referring to “faith-based organizations” and recipient organizations’ “faith-based character” might be read as intending that pervasively sectarian religious organizations be eligible to receive direct governmental funding. More recently, however, the term “faith-based organizations” has commonly been understood not to be limited to pervasively sectarian organizations, but also to encompass social-service organizations that are religiously affiliated and organizations with secular social service missions that are motivated or compelled by religious faith or obligation. In this memorandum, the term “faith-based organizations” is not limited to pervasively sectarian organizations.

² See Kendrick, 487 U.S. at 608-11 (holding that Adolescent Family Life Act grants may be awarded to religious institutions in light of the availability of such grants to a fairly “wide spectrum of public and private organizations” regardless of religious nature); see also, e.g., Roemer v. Board of Public Works, 426 U.S. 736 (1976) (plurality opinion) (upholding grant program for colleges and universities as applied to schools with religious affiliations).

³ In Kendrick, all nine Justices accepted the principle that government funding of religious activities would be impermissible. 487 U.S. at 611-12, 615, 621 (Establishment Clause would be violated if public monies were used to fund “‘indoctrination into the beliefs of a particular religious faith’” or to “‘advance the religious mission’ of the religious institution receiving aid.”) (quoting School Dist. of Grand Rapids v. Ball,

‘specifically religious activit[ies] in an otherwise substantially secular setting.’” Kendrick, 487 U.S. at 621 (quoting Hunt v. McNair, 413 U.S. 734, 743 (1973)).³ Accordingly, it would be impermissible for the government to permit direct federal financial assistance to be used for religious worship, religious instruction, or proselytizing. See Kendrick, 487 U.S. at 621 (constitutionality of a particular organization’s use of government grants would depend in part on whether the grantee “use[s] materials that have an explicitly religious content or are designed to inculcate the views of a particular religious faith”).⁴ For example, it may be the case that a religious organization effectively discourages teen drug use by conveying religious messages to teenagers, and by urging such teenagers to confirm, intensify, or rely upon their religious faith. Such an organization cannot use government funds to discourage teen drug use in this religious manner, even if such religious activities are effective in achieving the secular goal.⁵ In order to ensure that direct government funds are not used for sectarian purposes, it is helpful if the government provides some sort of “mechanism whereby the [government agency] can police the [aid] that [is] given out . . . to ensure that federal [aid is] not used for impermissible purposes.” Kendrick, 487 U.S. at 615.

2. The government may not provide financial aid directly to “pervasively sectarian” institutions, which the Court has defined as institutions in which “‘religion is so pervasive that a substantial portion of [their] functions are subsumed in the religious mission.’” Kendrick, 487 U.S. at 610 (quoting Hunt, 413 U.S. at 743).⁶ This second principle is ancillary to the first one. The Court presumes that any direct financial aid to a “pervasively sectarian” organization will

473 U.S. 373, 385 (1985)); id. at 623 (O’Connor, J., concurring) (“[A]ny use of public funds to promote religious doctrines violates the Establishment Clause.”); id. at 624 (Kennedy, J., concurring) (reasoning that the Establishment Clause would be violated if funds “are in fact being used to further religion”); id. at 634-48 (Blackmun, J., dissenting) (opining that government aid may not be used to advance religion, even if aid were intended for secular purposes).

⁴ By the same token, it is impermissible for the government to finance the construction or refinancing of a facility used or to be used for sectarian instruction or as a place for religious worship, Tilton v. Richardson, 403 U.S. 672, 683 (1971) (plurality opinion); see also Kendrick, 487 U.S. at 612 (aid may not be used to “‘advance the religious mission’ of the religious institution receiving aid”).

⁵ Consistent with this constitutional requirement, many federal statutes include restrictions against such religious uses of government funds. See, e.g., 42 U.S.C.A. § 9858k(a) (“No financial assistance [for child-care services and related activities] provided under this subchapter . . . shall be expended for any sectarian purpose or activity, including sectarian worship or instruction.”); see also Tilton, 403 U.S. at 675 (plurality opinion) (quoting statute that excluded from eligibility for construction and repair grants and loans “‘any facility used or to be used for sectarian instruction or as a place for religious worship’”).

⁶ See also id. at 621 (noting that plaintiffs could prevail under the Establishment Clause if they could show that aid flowed to grantees that were “pervasively sectarian religious institutions”); Columbia Union College v. Clarke, 159 F.3d 151, 157-62 (4th Cir. 1998), cert. denied, 119 S. Ct. 2357 (1999); Brief for the [Secretary of Health and Human Services] at 27-41, Bowen v. Kendrick, 487 U.S. 589 (1988) (Nos. 87-153, 87-431, 87-462); Memorandum for John J. Knapp, General Counsel, Department of Housing and Urban Development, from Theodore B. Olson, Assistant Attorney General, Office of Legal Counsel, Re: First Amendment Issues Implicated in Section 202 Loans and the Community Development Block Grant Program (July 1, 1983).

invariably support religious activity. Accordingly, statutes that authorize direct government aid to religious organizations should be construed as forbidding the funding of pervasively sectarian organizations and as permitting Federal, State, and local governments involved in disbursing funds to take into account the structure and operations of a religious organization in determining whether such an organization is pervasively sectarian. See, e.g., Statement of the President on Signing the Community Opportunities, Accountability, and Training and Educational Services Act of 1998, 34 Weekly Comp. Pres. Doc. 2148 (Oct. 27, 1998).⁷

Churches and church elementary and secondary schools are examples of organizations that traditionally have been deemed pervasively sectarian. Other religious organizations must be evaluated on a case-by-case basis. "The criteria for assessing whether and institution is pervasively sectarian are complex, elusive, and heavily fact intensive." Columbia Union College v. Clarke, 159 F.3d 151, 169 (4th Cir. 1998), cert. denied, 119 S. Ct. 2357 (1999). For example, courts will examine whether the organization has religious qualifications for admission and hiring, the degree to which the organization is institutionally autonomous from a particular religious denomination, and whether the organization has bylaws or policies that prohibit members or employees from deviating from religious doctrine.⁸ Government agencies should examine these and other relevant factors in determining whether an applicant for direct government aid is pervasively sectarian.⁹

3. Religious organizations that are not pervasively sectarian may receive

⁷ The law is more uncertain, and may be more permissive, with respect to certain forms of tangible aid, such as classroom materials and equipment. In a case currently pending in the Supreme Court, Mitchell v. Helms, Nos. 98-1648, 98-1671, cert. granted, 119 S. Ct. 2336 (1999), the United States has taken the position that the government can constitutionally provide certain forms of "[d]irect material aid" (other than funding) to religious primary and secondary schools (which the Court often has treated as pervasively sectarian institutions) so long as: the aid is allocated according to neutral, secular criteria; the aid does not supplant the school's otherwise available resources; the aid does not so substantially subsidize the school's secular functions that the school is enabled to shift its resources to sectarian functions; the government implements safeguards that ensure the aid is not used for sectarian purposes; and such safeguards do not involve public authorities in pervasive monitoring of the school's functions amounting to impermissible entanglement. The Court's decision in Mitchell might affect the constitutional questions addressed in this memorandum. We note, however, that if the Court were, in Mitchell, to adopt the United States's argument, such a holding would not necessarily authorize the direct provision of government funds to pervasively sectarian organizations.

⁸ See, e.g., Kendrick, 487 U.S. at 620 n.16; School Dist. of Grand Rapids v. Ball, 473 U.S. 373, 384 n.6 (1985); Roemer, 426 U.S. at 755-59 (plurality opinion); Hunt, 413 U.S. at 743-44; Committee for Public Educ. & Religious Liberty v. Nyquist, 413 U.S. 756, 767-68 (1973); Columbia Union College, 159 F.3d at 164-69.

⁹ The analysis in this section addresses only situations in which the government provides funds directly to organizations to enable such organizations to provide public services (including services to private individuals). This memorandum does not address statutes and programs pursuant to which the government instead provides aid directly to the ultimate individual beneficiaries of the program and then permits such persons to use the aid at organizations of their choosing. Cf. Zobrest v. Catalina Foothills School Dist., 509 U.S. 1, 9-10 (1993); Witters v. Washington Dept. of Servs. for the Blind, 474 U.S. 481, 487 (1986); Mueller v. Allen, 463 U.S. 388, 398-99 (1983).

government aid on the same terms as other, analogous community and nonprofit organizations, on the basis of neutral, secular criteria. But the government may not prefer a particular organization because it is, or is not, religious or religiously affiliated. It is a “principle at the heart of the Establishment Clause” that the government “should not prefer . . . religion to irreligion.” Board of Educ. of Kiryas Joel Village School Dist. v. Grumet, 512 U.S. 687, 703 (1994); see also, e.g., id. at 703-05; Kendrick, 487 U.S. at 607-09 (stressing the “neutrality” of the government aid to private organizations); Texas Monthly, Inc. v. Bullock, 489 U.S. 1 (1989);¹⁰ cf. Kiryas Joel, 512 U.S. at 709 (dicta) (Establishment Clause would prohibit a “patently unequal law” that provided school transportation only for children attending Christian day schools).¹¹ Accordingly, no government aid program should be limited to religious or “faith-based” organizations, and the government must not otherwise prefer such organizations over others, e.g., by setting aside a particular portion of funds for them. Similarly, nonpervasively sectarian religious organizations should not be disfavored for aid because of their religious character or affiliations. The criteria for funding should be neutral and secular, and the government must not “convey[] or attempt[] to convey a message that religion or a particular religious belief is favored or preferred.” Wallace v. Jaffree, 472 U.S. 38, 70 (1985) (O’Connor, J., concurring in the judgment). The government may support grants to religious organizations because of the organizations’ effectiveness in helping to address social problems through secular means, but not because of governmental support or preference for such organizations’ religious nature, activities or affiliations.

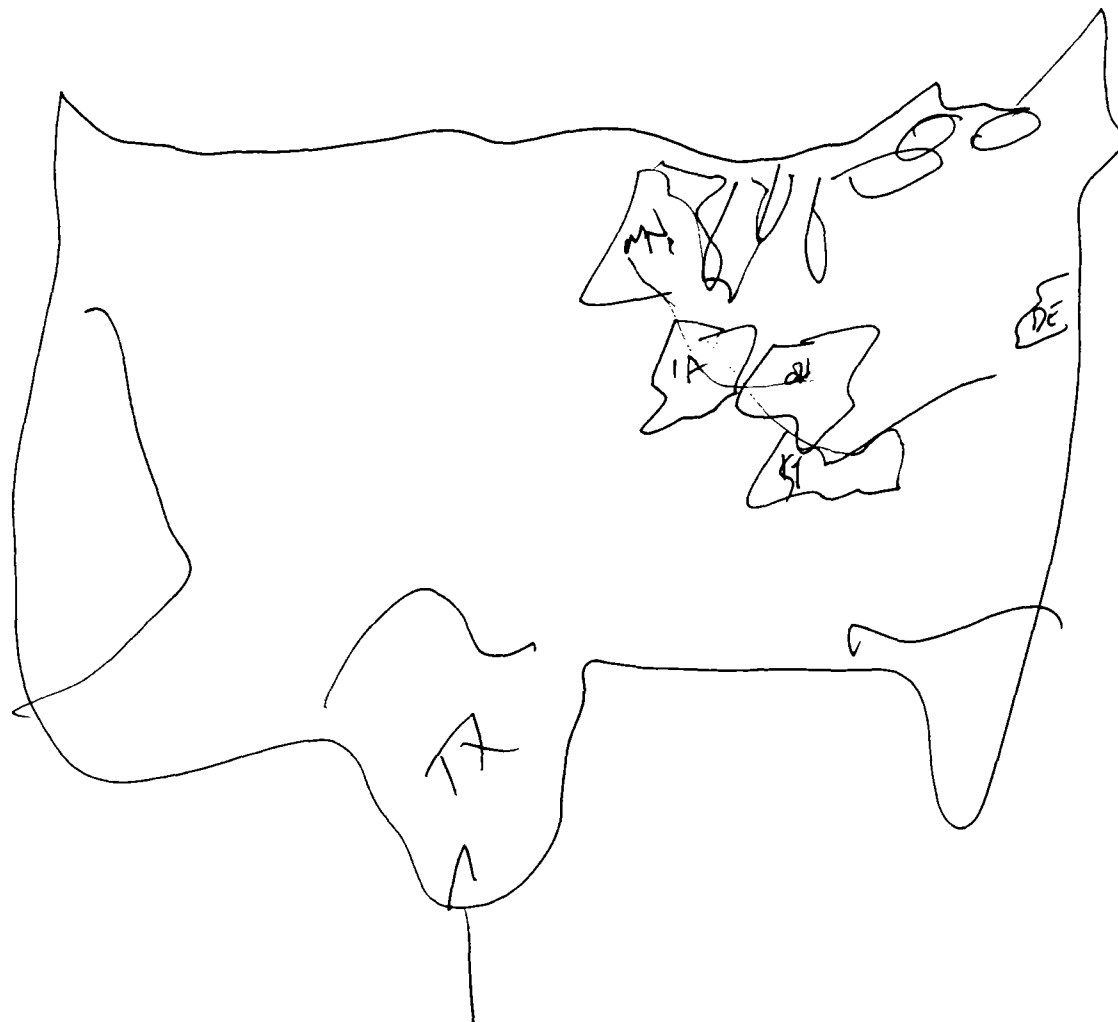
4. The government may not prefer certain religious sects or organizations over others, except on the basis of secular criteria unrelated to the organizations’ religious affiliations or beliefs. “The clearest command of the Establishment Clause is that one religious denomination cannot be officially preferred over another.” Larson v. Valente, 456 U.S. 228,

¹⁰ In Texas Monthly, for example, the Court invalidated a sales-tax exemption provided exclusively to religious periodicals. The plurality reasoned that “when government directs a subsidy exclusively to religious organizations that is not required by the Free Exercise Clause and that either burdens nonbeneficiaries markedly or cannot reasonably be seen as removing a significant state-imposed deterrent to the free exercise of religion, . . . it ‘provide[s] unjustifiable awards of assistance to religious organizations’ and cannot but ‘conve[y] a message of endorsement’ to slighted members of the community.” 489 U.S. at 15 (quoting Corporation of Presiding Bishop of Church of Jesus Christ of Latter-day Saints v. Amos, 483 U.S. 327, 348 (1987) (O’Connor, J., concurring in the judgment)). A government may, of course, provide such benefits to religious adherents when such persons are included in a broader class of similarly situated beneficiaries, defined in secular terms and reflecting some legitimate secular aim. See Texas Monthly, 489 U.S. at 15-17 (plurality opinion) (scope of beneficiary class must reflect legitimate “secular aim,” and not have the “purpose or effect of sponsoring certain religious tenets or religious belief in general”).

¹¹ See also Mueller, 463 U.S. at 396-97 (in upholding tax deduction for educational expenses against Establishment Clause challenge, the Court deemed it “particularly significant” that “the deduction is available for educational expenses incurred by all parents, including those whose children attend public schools and those whose children attend nonsectarian private schools or sectarian private schools”); Nyquist, 413 U.S. at 788-89.

244 (1982); see also Kiryas Joel, 512 U.S. at 706-07 (“whatever the limits of permissible legislative accommodations may be, . . . it is clear that neutrality as among religions must be honored”).¹² Moreover, facially neutral criteria may run afoul of this principle if they “effectively distinguish[]” among religions, Larson, 456 U.S. at 247 n.23, or are “drafted with the explicit intention” of preferring “particular religious denominations,” id. at 254-55 (quoting Gillette v. United States, 401 U.S. 437, 452 (1971)). This principle of denominational neutrality also generally applies to forms of government aid other than direct funding. Accordingly, the criteria for government “partnerships” with private organizations must be neutral and secular.

¹² A denominational preference is treated as suspect, and must be invalidated unless it is justified by a compelling governmental interest, and unless it is closely fitted to further that interest. Larson, 456 U.S. at 246-47.



Office of Senator Max Cleland
United States Senate
461 Dirksen Senate Office Building
Washington, DC 20510
Phone: (202) 224-3521
Fax: (202) 224-0072

FAX COVER SHEET

FAX NUMBER TRANSMITTED TO: _____ PHONE #: _____

To: Andy Rotherham

Organization: _____

From: Lynne Kennerly (224-5220)

Date: 4/26/2000

Number of Pages: 15 (Including this cover sheet)

COMMENTS:

If you do not receive all the pages of this transmission,
please contact (202) 224-3521

NOTE: THIS DOCUMENT MAY CONTAIN PRIVILEGED AND NONPUBLIC INFORMATION. IT IS INTENDED ONLY FOR THE USE OF THE INDIVIDUAL OR ENTITY NAMED ABOVE, AND OTHERS WHO SPECIFICALLY HAVE BEEN AUTHORIZED TO RECEIVE IT.

MAX CLELAND
GEORGIA

Telephone: (202) 224-3821
TDD/TTY: (202) 224-3283
www.senate.gov/~cleland

COMMITTEES:
ARMED SERVICES
COMMERCE
GOVERNMENTAL AFFAIRS
SMALL BUSINESS

United States Senate

WASHINGTON, DC 20510-1005

April 25, 2000

IMMIGRANTS TO NEW AMERICANS ACT

Dear Colleague:

From Dallas to Duluth to Denver, America's communities are facing an increasing influx of immigrant and refugee families. In 1997, there were an estimated 25.8 million foreign-born people residing in the United States. This is the largest number of foreign-born individuals living here in U.S. history and represents a 30% increase over the 1990 census figure of 19.8 million people. The U.S. Census Bureau is estimating that these recently arrived immigrants and refugees will account for 75% of the U.S. population growth over the next 50 years!

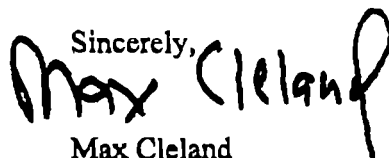
The changing demographics are impacting school systems not only in our major metropolitan areas, but also rural areas like Albertville, Alabama, where the number of Spanish-speaking students went from eight in 1992 to 450 in 1999. In Waterloo, Iowa there are 400 Bosnian children who recently arrived without knowing our language, culture or customs. The A. C. Kiefer School in Wausau, Wisconsin, which is 58% Asian, has had recent arrivals from the Hmong, Laotian and other Asian cultures. In Maine's rural communities, the arrival of thousands of immigrant Mexican families has overwhelmed the schools and community-based organizations that provide them with services. Recent Census Bureau estimates are showing that between 1990 and 1998 the percentage of foreign-born residents in 11 southern states has grown by double digits. In my state of Georgia, several communities have witnessed their Hispanic population soar over the last decade. In Dalton, Georgia's public schools, for example, enrollment of Hispanic students is now 47 percent, up from just 4 percent ten years ago. These examples are illustrative of what is happening to student and community populations nationwide, and they underscore the need to provide the services and support that will enable the new arrivals to our country to successfully participate in America's schools and communities.

For this reason, I will soon be introducing a bill which would establish a competitive grant program within the Department of Education to assist schools and communities that are experiencing an infusion of recently arrived foreign-born families. Specifically, this grant program would provide funding to partnerships of local school districts and community-based organizations to develop model programs that assist culturally and linguistically diverse children achieve success in America's schools and provide their families with access to comprehensive community services, including health care, child care, job training and transportation.

The Immigrants to New Americans Act is endorsed by the National Association for Bilingual Education, the League of United Latin American Citizens, the India Abroad Center for Political Awareness, the National Korean American Service and Education Consortium, and the National Council of La Raza.

If you are interested in cosponsoring this legislation or have any questions about it, please contact Lynn Kimmerly of my staff at 224-5220.

Sincerely,



Max Cleland

SUITE 1700
75 SPRING STREET, S.W.
ATLANTA, GA 30303-3309
(404) 331-4811

SUITE 101
235 ROOSEVELT AVENUE
ALBANY, GA 31701-2372
(912) 436-7796

611 TOLSON STREET
AUGUSTA, GA 30901-2324
(706) 722-4040

SUITE 101
120 12TH STREET
COLUMBUS, GA 31902-2481
(706) 648-7706

203 MARTIN LUTHER KING JR. BLVD.
DALTON, GA 30721-3004
(706) 276-8906

SUITE 203
401 CHERRY STREET
MACON, GA 31201-3384
(912) 766-1779

SUITE A
440 MALL BOULEVARD
SAVANNAH, GA 31406-4823
(812) 362-8283

106TH CONGRESS
2D SESSION

S. _____

IN THE SENATE OF THE UNITED STATES

Mr. CLELAND introduced the following bill; which was read twice and referred
to the Committee on _____

A BILL

To ensure that immigrant students and their families receive
the services that the students and families need to suc-
cessfully participate in elementary schools, secondary
schools, and communities, in the United States, and for
other purposes.

1 *Be it enacted by the Senate and House of Representa-*
2 *tives of the United States of America in Congress assembled,*

3 **SECTION 1. IMMIGRANTS TO NEW AMERICANS MODEL PRO-**
4 **GRAMS.**

5 (a) **SHORT TITLE.**—This section may be cited as the
6 “Immigrants to New Americans Act”.

7 (b) **FINDINGS.**—Congress finds the following:

1 (1) In 1997, there were an estimated
2 25,800,000 foreign-born individuals residing in the
3 United States. That number is the largest number
4 of such foreign-born individuals ever in United
5 States history and represents a 6,000,000, or 30
6 percent, increase over the 1990 census figure of
7 19,800,000 of such foreign-born individuals. The
8 Bureau of the Census estimates that the recently ar-
9 rived immigrant population (including the refugee
10 population) currently residing in the Nation will ac-
11 count for 75 percent of the population growth in the
12 United States over the next 50 years.

13 (2) For millions of immigrants settling into the
14 Nation's hamlets, towns, and cities, the dream of
15 "life, liberty, and the pursuit of happiness" has be-
16 come a reality. The wave of immigrants, from var-
17 ious nationalities, who have chosen the United
18 States as their home, has positively influenced the
19 Nation's image and relationship with other nations.
20 The diverse cultural heritage of the Nation's immi-
21 grants has helped define the Nation's culture, cus-
22 toms, economy, and communities. By better under-
23 standing the people who have immigrated to the Na-
24 tion, individuals in the United States better under-
25 stand what it means to be an American.

1 (3) There is a critical shortage of teachers with
2 the skills needed to educate immigrant students and
3 their families in nonconcentrated, nontraditional, im-
4 migrant communities as well as communities with
5 large immigrant populations. The large influx of im-
6 migrant families over the last decade presents a na-
7 tional dilemma: The number of such families with
8 school-age children, requiring assistance to success-
9 fully participate in elementary schools, secondary
10 schools, and communities in the United States, is in-
11 creasing without a corresponding increase in the
12 number of teachers with skills to accommodate their
13 needs.

14 (4) Immigrants arriving in communities across
15 the Nation generally settle into high-poverty areas,
16 where funding for programs to provide immigrant
17 students and their families with the services the stu-
18 dents and families need to successfully participate in
19 elementary schools, secondary schools, and commu-
20 nities in the United States is inadequate.

21 (5) The influx of immigrant families settling
22 into many United States communities is often the
23 result of concerted efforts by local employers who
24 value immigrant labor. Those employers realize that
25 helping immigrants to become productive, pros-

1 perous members of a community is beneficial for the
2 local businesses involved, the immigrants, and the
3 community. Further, local businesses benefit from
4 the presence of the immigrant families because the
5 families present businesses with a committed and ef-
6 fective workforce and help to open up new market
7 opportunities. However, many of the communities
8 into which the immigrants have settled need assist-
9 ance in order to give immigrant students and their
10 families the services the students and families need
11 to successfully participate in elementary schools, sec-
12 ondary schools, and communities, in the United
13 States.

14 (c) PURPOSE.—The purpose of this section is to es-
15 tablish a grant program, within the Department of Edu-
16 cation, that provides funding to partnerships of local edu-
17 cational agencies and community-based organizations for
18 the development of model programs to provide to immi-
19 grant students and their families the services the students
20 and families need to successfully participate in elementary
21 schools, secondary schools, and communities, in the
22 United States.

23 (d) DEFINITIONS.—In this section:

24 (1) COMMUNITY-BASED ORGANIZATION; ELE-
25 MENTARY SCHOOL; LOCAL EDUCATIONAL AGENCY;

1 SECONDARY SCHOOL.—The terms “community-based
2 organization”, “elementary school”, “local edu-
3 cational agency”, and “secondary school” have the
4 meanings given the terms in section 14101 of the
5 Elementary and Secondary Education Act of 1965
6 (20 U.S.C. 8801).

7 (2) IMMIGRANT.—The term “immigrant” has
8 the meaning given the term in section 101 of the
9 Immigration and Nationality Act (8 U.S.C. 1101).

10 (3) SECRETARY.—The term “Secretary” means
11 the Secretary of Education.

12 (e) PROGRAM AUTHORIZED.—

13 (1) IN GENERAL.—The Secretary is authorized
14 to award not more than 10 grants in a fiscal year
15 to eligible partnerships for the design and implemen-
16 tation of model programs to—

17 (A) assist immigrant students to achieve in
18 elementary schools and secondary schools in the
19 United States by offering such educational serv-
20 ices as English as a second language classes,
21 literacy programs, programs for introduction to
22 the education system, and civics education; and

23 (B) assist parents of immigrant students
24 by offering such services as parent education
25 and literacy development services and by coordi-

1 nating activities with other entities to provide
2 comprehensive community social services such
3 as health care, job training, child care, and
4 transportation services.

5 (2) DURATION.—Each grant awarded under
6 this section shall be awarded for a period of not
7 more than 5 years. A partnership may use funds
8 made available through the grant for not more than
9 1 year for planning and program design.

10 (f) APPLICATIONS FOR GRANTS.—

11 (1) IN GENERAL.—Each eligible partnership de-
12 siring a grant under this section shall submit an ap-
13 plication to the Secretary at such time and in such
14 manner as the Secretary may require.

15 (2) ELIGIBLE PARTNERSHIPS.—To be eligible
16 to receive a grant under this section, a
17 partnership—

18 (A) shall include—

19 (i) at least 1 local educational agency;

20 and

21 (ii) at least 1 community-based orga-
22 nization; and

23 (B) may include another entity such as an
24 institution of higher education, a local or State
25 government agency, a private sector entity, or

1 another entity with expertise in working with
2 immigrants.

3 (3) REQUIRED DOCUMENTATION.—Each appli-
4 cation submitted by a partnership under this section
5 for a proposed program shall include documentation
6 that—

7 (A) the partnership has the qualified per-
8 sonnel required to develop, administer, and im-
9 plement the proposed program; and

10 (B) the leadership of each participating
11 school has been involved in the development and
12 planning of the program in the school.

13 (4) OTHER APPLICATION CONTENTS.—Each ap-
14 plication submitted by a partnership under this sec-
15 tion for a proposed program shall include—

16 (A) a list of the organizations entering into
17 the partnership;

18 (B) a description of the need for the pro-
19 posed program, including data on the number
20 of immigrant students, and the number of such
21 students with limited English proficiency, in the
22 schools or school districts to be served through
23 the program and the characteristics of the stu-
24 dents described in this subparagraph,
25 including—

1 (i) the native languages of the stu-
2 dents to be served;

3 (ii) the proficiency of the students in
4 English and the native languages;

5 (iii) achievement data for the students
6 in—

7 (I) reading or language arts (in
8 English and in the native languages,
9 if applicable); and

10 (II) mathematics; and

11 (iv) the previous schooling experiences
12 of the students;

13 (C) a description of the goals of the pro-
14 gram;

15 (D) a description of how the funds made
16 available through the grant will be used to sup-
17 plement the basic services provided to the immi-
18 grant students to be served;

19 (E) a description of activities that will be
20 pursued by the partnership through the pro-
21 gram, including a description of—

22 (i) how parents, students, and other
23 members of the community, including
24 members of private organizations and non-

1 profit organizations, will be involved in the
2 design and implementation of the program;

3 (ii) how the activities will further the
4 academic achievement of immigrant stu-
5 dents served through the program;

6 (iii) methods of teacher training and
7 parent education that will be used or devel-
8 oped through the program, including the
9 dissemination of information to immigrant
10 parents, that is easily understandable in
11 the language of the parents, about edu-
12 cational programs and the rights of the
13 parents to participate in educational deci-
14 sions involving their children; and

15 (iv) methods of coordinating com-
16 prehensive community social services to as-
17 sist immigrant families;

18 (F) a description of how the partnership
19 will evaluate the progress of the partnership in
20 achieving the goals of the program;

21 (G) a description of how the local edu-
22 cational agency will disseminate information on
23 model programs, materials, and other informa-
24 tion developed under this section that the local
25 educational agency determines to be appro-

1 priate for use by other local educational agen-
2 cies in establishing similar programs to facili-
3 tate the educational achievement of immigrant
4 students;

5 (H) an assurance that the partnership will
6 annually provide to the Secretary such informa-
7 tion as may be required to determine the effec-
8 tiveness of the program; and

9 (I) any other information that the Sec-
10 retary may require.

11 (g) SELECTION OF GRANTEES.—

12 (1) CRITERIA.—The Secretary, through a peer
13 review process, shall select partnerships to receive
14 grants under this section on the basis of the quality
15 of the programs proposed in the applications sub-
16 mitted under subsection (f), taking into consider-
17 ation such factors as—

18 (A) the extent to which the program pro-
19 posed in such an application effectively address-
20 es differences in language, culture, and cus-
21 toms;

22 (B) the quality of the activities proposed
23 by a partnership;

24 (C) the extent of parental, student, and
25 community involvement;

1 (D) the extent to which comprehensive
2 community social services are made available;

3 (E) the quality of the plan for measuring
4 and assessing success; and

5 (F) the likelihood that the goals of the pro-
6 gram will be achieved.

7 (2) GEOGRAPHIC DISTRIBUTION OF PRO-
8 GRAMS.—The Secretary shall approve applications
9 under this section in a manner that ensures, to the
10 extent practicable, that programs assisted under this
11 section serve different areas of the Nation, including
12 urban, suburban, and rural areas, with special atten-
13 tion to areas that are experiencing an influx of im-
14 migrant groups (including refugee groups), and that
15 have limited prior experience in serving the immi-
16 grant community.

17 (h) EVALUATION AND PROGRAM DEVELOPMENT.—

18 (1) REQUIREMENT.—Each partnership receiv-
19 ing a grant under this section shall—

20 (A) conduct a comprehensive evaluation of
21 the program assisted under this section, includ-
22 ing an evaluation of the impact of the program
23 on students, teachers, administrators, parents,
24 and others; and

1 (B) prepare and submit to the Secretary a
2 report containing the results of the evaluation.

3 (2) EVALUATION REPORT COMPONENTS.—Each
4 evaluation report submitted under this section for a
5 program shall include—

6 (A) data on the partnership's progress in
7 achieving the goals of the program;

8 (B) data showing the extent to which all
9 students served by the program are meeting the
10 State's student performance standards,
11 including—

12 (i) data comparing the students
13 served to other students, with regard to
14 grade retention and academic achievement
15 in reading and language arts, in English
16 and in the native languages of the students
17 if the program develops native language
18 proficiency, and in mathematics; and

19 (ii) a description of how the activities
20 carried out through the program are co-
21 ordinated and integrated with the overall
22 school program of the school in which the
23 program described in this section is carried
24 out, and with other Federal, State, or local

1 programs serving limited English proficient
2 students;

3 (C) data showing the extent to which fami-
4 lies served by the program have been afforded
5 access to comprehensive community social serv-
6 ices; and

7 (D) such other information as the Sec-
8 retary may require.

9 (i) ADMINISTRATIVE FUNDS.—A partnership that re-
10 ceives a grant under this section may use not more than
11 5 percent of the grant funds received under this section
12 for administrative purposes.

13 (j) AUTHORIZATION OF APPROPRIATIONS.—For the
14 purpose of carrying out this section, there are authorized
15 to be appropriated \$10,000,000 for fiscal year 2001 and
16 such sums as may be necessary for each of the 4 suc-
17 ceeding fiscal years.