DOCUMENT RESUME

ED 388 590 SO 025 532

TITLE	The 1994 NAEP Achievement Levels-Setting Process for U.S. History and Geography. Preliminary Report.
INSTITUTION	American Coll. Testing Program, Iowa City, Iowa.
SPONS AGENCY	National Assessment Governing Board, Washington, DC.
PUB DATE	May 95
NOTE	101p.
AVAILABLE FROM	American College Testing, P.O. Box 168, Iowa City, IA 52243-0168.
PUB TYPE	Information Analyses (070) -- Reports - Descriptive (141)
EDRS PRICE	MF01/PC05 Plus Postage.
DESCRIPTORS	*Academic Achievement; *Educational Assessment; Elementary Secondary Education; *Evaluation Methods; *Geography; Grade 4; Grade 8; Grade 12; National Programs; *Student Evaluation; *United States History
IDENTIFIERS	National Assessment of Educational Progress

ABSTRACT

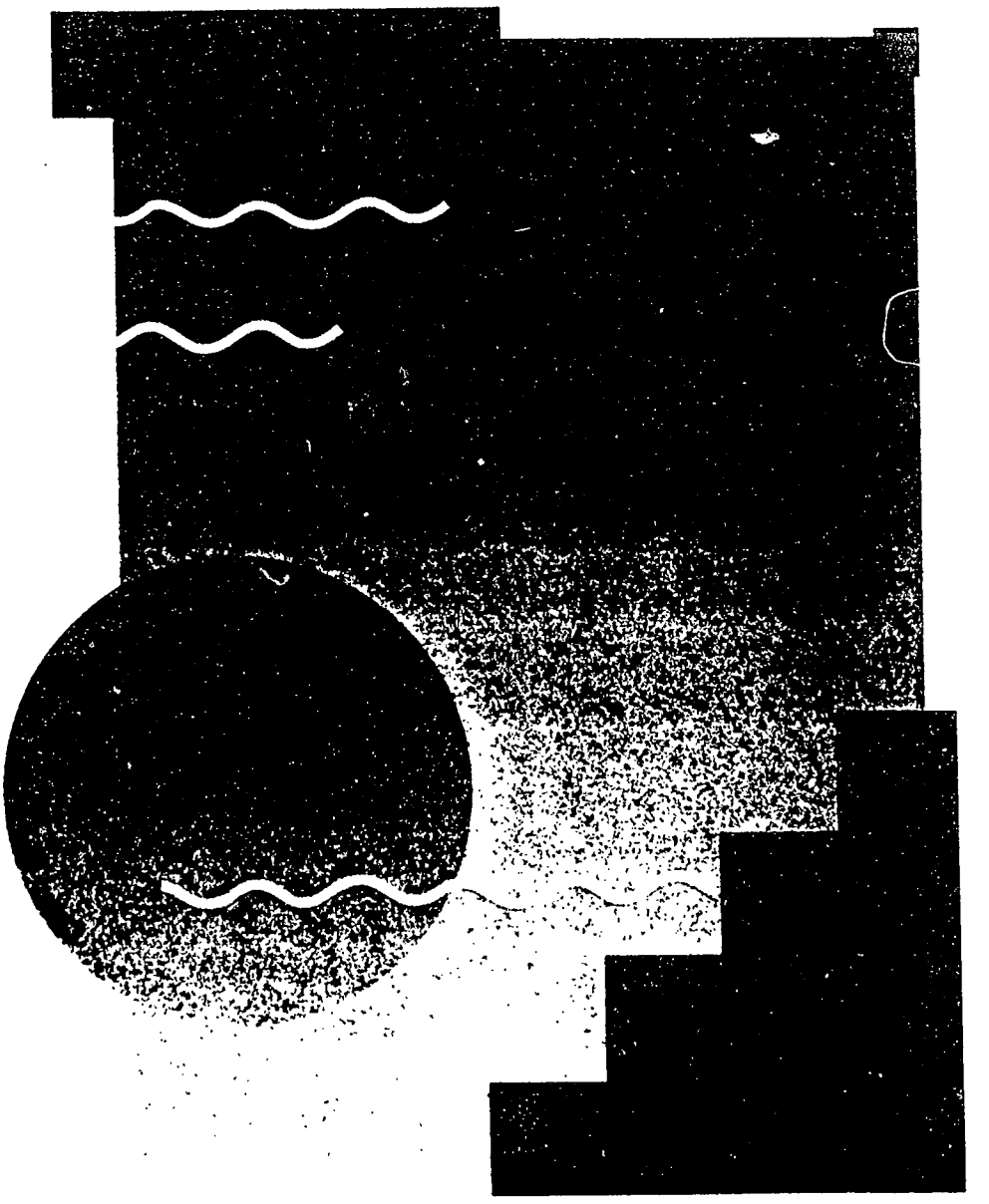
This report describes the process of establishing the National Assessment of Educational Progress (NAEP) Achievement Levels-Setting (ALS) to formulate expectations for the geography and U.S. history assessments in 1994 for grades 4, 8, and 12. The volume includes the following: (1) "List of Tables"; (2) "List of Figures"; (3) "Preface"; (4) "Overview"; (5) "Introduction"; (6) "Panelists"; (7) "Achievement Levels-Setting Process"; (8) "Achievement Level Cutscores"; (9) "Achievement Levels Descriptions"; (10) "Exemplar Items"; (11) "Panelist Evaluations" which include understanding of achievement levels descriptions, conceptions of borderline performance, rating methods, and difference in performance estimates by item type; (12) "Consequences Data"; (13) "Conclusions", and (14) "Validation Research Studies" which is subdivided into similarity research studies and booklet classification study. (EH)

ACT

ED 388 590

Preliminary Report on the 1994 NAEP Achievement Levels-Setting Process for U.S. History and Geography

Presented by American College Testing
May 1995



Preliminary Report on the 1994 NAEP Achievement Levels-Setting Process for U.S. History and Geography

May 1995

Table of Contents

List of Tables

List of Figures

Preface

This report has been prepared, in part, to inform interested individuals, groups, and organizations regarding the Achievement Levels-Setting Project for the 1994 National Assessment of Educational Progress in Geography and U.S. History. The Commissioner of Education Statistics is authorized to "release" the data. Until the Governing Board has made their final decisions and these decisions have been made public by duly authorized persons, the achievement levels data are limited to those included in this report. Every effort has been made to provide informative data for review and comment. Cautions are given regarding the interpretation of some data, however.

Input from various individuals, groups, and organizations is sought at this time because recommendations have not yet been submitted to NAGB for their final review and decision. Comment from the public will be recorded and presented to NAGB as part of the entire package of information to inform their decision.

Questions and comments are encouraged and appreciated. Written comments may be addressed to ACT and they will be forwarded, along with other recommendations to NAGB.

Susan Cooper Loomis, Ph.D.
NAEP ALS Project Director
May 27, 1995

Overview

In October 1993, the National Assessment Governing Board (NAGB) awarded a contract to set achievement levels on the 1994 National Assessment of Educational Progress (NAEP) in Geography and U.S. History and the 1996 Science NAEP to American College Testing (ACT).

The 1994 achievement levels-setting (ALS) process consisted of several steps. The planning and design of the process drew heavily from the experiences of the 1992 ALS process and evaluations of that process. Comments and suggestions solicited from various individuals, groups, and organizations during the design stages contributed to the final features of the process design.

The planning and design stages culminated in the pilot studies for both geography and U.S history. The pilot studies tested all procedures under consideration for implementation in the achievement levels-setting process. In addition, several research studies were conducted. Among the most important research components was the study of rating methods and the differences in cutscores resulting from the different methods. Findings from those studies led to the choice of methods implemented in the ALS process.

Other research issues of particular interest related to the use of more "holistic" rating methods and information regarding student performance relative to the achievement levels. These methods were called for by the National Academy of Education report.[1]

Following the pilot studies and the evaluation of the rating and evaluation data from those studies, the ALS panels were convened. The ALS panels for geography were convened November 12-16, 1995, and the U.S. history ALS panels were convened December 2-6, 1993. The results of these ALS meetings are summarized in this report.

In addition to the design and implementation of the ALS process, however, NAGB requested that ACT develop additional research studies to examine the levels recommended by the ALS panelists. Specifically, the request was that studies be designed to address the validity of the ALS process and the levels set.

ACT proposed several studies. The Technical Advisory Committee on Standard Setting (TACSS) reviewed these proposed studies and recommended three for implementation. Priorities were assigned to the three studies, and the ***Similarity Classification Study*** was given top priority, followed by the ***Booklet Classification Study***, and a ***NAEP Consumer Survey***. Very general and preliminary results from the first two studies are presented in this report. The survey has not yet been conducted. More complete information on each of these studies is available upon request.

---

[1] National Academy of Education Panel, **Setting Performance Standards for Student Achievement**, (Stanford University, Stanford, CA.: Author, 1993)

Preliminary Report on the 1994 NAEP Achievement Levels-Setting Process for U.S. History and Geography

Introduction

The 1994 NAEP Achievement Levels-Setting (ALS) process was implemented November 12-16, 1994 for Geography and December 2-6, 1994 for U.S. History. An item-by-item rating methodology was used by panelists to rate the items in the 1994 item pool for each grade level tested by NAEP—Grades 4, 8, and 12.

Thirty panelists at each grade level were divided into item rating groups and items were divided into item rating pools. Half the panelists at each grade level rated half the items in the item pool for that grade. Panelists were divided into the two groups so that panelist type (teachers, other educators, and general public), race/ethnicity, sex, and region were represented as equally as possible in the two groups. Similarly, item blocks (units of approximately 17 items) were divided into two groups so that item characteristics such as difficulty (probability of correct response), content type, and item format (multiple choice, short answer, extended response) were as equivalent as possible.

The modified Angoff method[2] was used to rate multiple choice items and the mean estimation method[3] was used to rate all other items. Panelists were engaged in training exercises for approximately three days prior to the first of three item rating sessions. The training was designed to assure that panelists understood the contents of the assessment framework and shared a common understanding of the definitions of achievement at each level, to help panelists reach a common agreement on "borderline performance," to become familiar with the scoring rubrics and their use for each of the items in the grade-level assessment, and to give panelists a "reality check" with respect to student performance on constructed response items.

Panelists completed a process evaluation at the end of each day or after completing a stage in the process. Those evaluations were reviewed each day to monitor any potential problems or difficulties for panelists. Responses to some of the evaluation items are included in this report.

Following the final round of ratings, panelists were provided "consequences data" that showed estimates of the percentages of students scoring at or above each level. A questionnaire was designed to collect opinions of panelists regarding the correspondence between these estimates and their own expectations as to the consequences of the levels they had set.

---

[2] The procedure panelists were instructed to follow: think of the description of what students should know and be able to do at the (Basic/Proficient/Advanced) achievement level; focus on the minimum level of performance required to reach that level of achievement; estimate the percentage of students who are at the lower borderline of the achievement level who would answer the item correctly.

[3] To rate essay items (short answer or extended response items), panelists were instructed to follow the same procedure outlined for the modified Angoff ratings on multiple choice items except they estimated the average score on each item for students at the lower borderline.

Panelists

The panelist selection process is described more completely in the ***Design Document.***[4] Principles of sampling were used for drawing stratified random samples of school districts from a national database. The samples were stratified to represent the four NAEP regions approximately equally, districts with enrollments of at least 50,000 students, and districts with at least 25% of the population below the poverty level. A total of six samples were drawn for identifying ALS nominators: one sample each for U.S. history and geography from which nominators of panelists of the three types (teacher, other educators, and general public representatives) were identified.

Nominators of teachers were district superintendents, leaders of teacher organizations, or head persons of private schools. In addition, state social studies curriculum directors were invited to nominate teachers from any district within their state.[5] Nominators of the other educators included various persons who were career educators but not themselves K-12 classroom teachers. Nominators of the general public included the mayor (or other chief elected official of the primary city in the school district), the school board president, and the chair of the education committee of the local Chamber of Commerce.

Organizations such as the National Council for History Education, National Council for Geography Education, and the Geography Alliance were invited to contact nominators and recommend panelists. The Geography Alliance participated extensively in this process[6].

A total of 221 persons were nominated as geography panelists and 225 as history panelists to fill 90 positions on the panel for each NAEP subject. Panelists were selected to serve at a specific grade level, although they were occasionally nominated at more than one grade level. Panelists were nominated to represent a specific perspective: grade-level classroom teacher, other educators, and general public. Table 1a shows the distribution of the nominee pool for geography and Table 1b shows the distribution for the U.S. history nominee pool.

ACT project staff developed a computerized process for selecting panelists from the pool of nominees. Panelists were rated according to their qualifications (information provided on the nomination form). Panelists with the highest qualifications ratings had the highest probability of being selected, other things being equal. The selection program was set to yield panels with 55% of the members representing grade-level classroom teachers, 15% representing other (nonteacher) educators, and 30% representing the general public. The

---

[4] American College Testing, ***Design Document: Setting Achievement Levels on the 1994 National Assessment of Educational Progress in Geography and in U.S. History and the 1996 National Assessment of Educational Progress in Science*** (Iowa City, IA: Author, 1993).

[5] Only social studies curriculum directors from states for which one or more school districts had been drawn in the "teacher nominator sample" were invited to nominate teachers in their state as panelists. Unlike other nominators of teachers, these persons could nominate teachers from the state "at large," rather from specific districts drawn in the sample.

[6] The Geography Alliance state coordinators were sent letters to inform them of the persons who were identified as nominators in their state. The state coordinators were urged to contact the nominator and provide suggestions of "outstanding teachers" to be nominated as panelists.

distribution by panelist type was only one requirement to be met, however. The selection was set to yield panels with 20% "minority" racial/ethnic representation, up to 50% males, and regional representation approximately equal to the proportion of the U.S. population in each of the four NAEP regions.

Some of the persons who were selected could not serve at the time required. Data for the actual panelists who participated in the ALS process are presented in Tables 2a and 2b for geography and U.S. history, respectively. Summary information about "outstanding" credentials for panelists are included in Figure 1a for geography and Figure 1b for U.S. history.

Achievement Levels-Setting Process

Panelists selected through the previously described process were convened in St. Louis, Missouri for a five-day ALS process. The same agenda was followed for both geography and U.S. history. During the opening sessions, panelists were provided a complete overview, *via* a computerized presentation, of the process and the procedures to be followed. Each panelist was given a *Briefing Booklet* that described each task to be performed during each session, the purpose for the task, and how to perform the task.[7] In addition, a "glossary" of terms was included to help panelists with new and jargon terms.

During the first day, a presentation by NAGB staff described the NAEP program, including an explanation of the government agencies involved, the contractors, and the interrelationships of each. That presentation traced the development of the NAEP for the specific subject. During the first day, panelists were administered the NAEP exam and allowed to self-score the exam.

Each grade-level panel was led by a process facilitator and a content facilitator. With the exception of the fourth grade U.S. history content facilitator, all staff participated in both the pilot studies as well as the ALS meetings. Content facilitators were selected from the most highly recommended members of the Framework Panel for each subject. Process facilitators were trained extensively both before the pilot studies and before the ALS meetings. Content facilitators were included in at least one training session with process facilitators prior to the pilot study.

Content facilitators provided training in the Frameworks and the preliminary achievement levels descriptions. They took the lead in all sessions to review achievement levels descriptions. Training in the Framework and preliminary descriptions was concentrated during the first two days of the process. Panelists were engaged in exercises designed to give them experience in working with the achievement levels descriptions. A session was scheduled to discuss the achievement levels descriptions, and an opportunity was provided for panelists to modify the preliminary descriptions before each round of ratings.

---

[7] The ***Briefing Booklet*** for geography was distributed at the start of the ALS meetings. For history, it was distributed as part of the advanced materials sent to each panelist approximately two weeks prior to the meetings. Other advanced materials sent to all panelists included a ***Summary Design Document*** that had been recommended as a "user friendly description of the proce..." during public comment sessions held during the spring of 1994.

The exercises and training provided to panelists were designed to help them become more familiar with the items in the NAEP for their grade level and the scoring rubrics for the items, as well as to become familiar with the Frameworks and descriptions. Prior to the first round of ratings, panelists were asked to review student responses written to all essay/constructed response questions. This exercise was especially designed to provide panelists with a "reality" check on student performance, to allow them to become more familiar with the scoring rubrics for these items, and to help them form a concept of borderline performance for each achievement level.

Instructions for each exercise and task were provided in general sessions so that all panelists were given the same instructions. Process facilitators repeated the instructions and answered questions from panelists in the grade-level sessions. All tasks were performed by panelists in grade-level sessions.

Three rounds of item ratings were collected. Each panelist rated each item in his/her item pool (half the items for the grade). The item ratings were entered and analyzed on site. These item ratings were used to produce feedback information for panelists regarding their ratings relative to those of others in their grade, information about overall student performance on each item, information about student performance on the test booklet form that panelists had been administered at the start, and information about the consistency of their ratings on specific items.

The feedback was presented and panelists were instructed in the meaning of the feedback and how to use it in subsequent rounds of ratings. Time was provided for panelists to ask questions and discuss the feedback before beginning another round of ratings. Feedback from the previous round was updated after each round of ratings and presented to panelists along with new information. Following Round 3, all feedback was again updated relative to the ratings for this, the final round.

Achievement Level Cutscores

The cutscores in Tables 3a (geography) and 3b (U.S. history) are reported on the ACT NAEP-like scale. This scale is linearly related to the NAEP scale, which means that the numerical levels and relative differences observed on this scale are proportional to values on the NAEP scale. The approximate proportion of score points required for students at each achievement level is included in Tables 3a and 3b as well. These data are provided ***only*** because actual performance data cannot be released at this time. These data reporting the proportion of possible points for each achievement level are peculiar to the items on this exam, the sample of schools, and the students assessed. If the assessments were administered again, the percentages would change. These proportions cannot be compared from one administration to another. Nonetheless, they do provide *some* information regarding the relative levels in a metric that is more easily understood than the "ACT NAEP-like" scale.

The "percent of possible points correct" data show that students needed to get at least one-third of the possible points for performance at the Basic level. The percentage increases across each grade for each subject. At the Proficient level, students needed to get nearly two-thirds of the possible points. Finally, at the Advanced level, students needed more than three-quarters of the total possible points on the items included on the assessment for the grade.

In order to get a better understanding of what these data mean, it is helpful to compare the percent of possible points at each cutscores to the percent of possible points scored by "the average" student who took the NAEP at each grade level. The percent of possible points at the mean for each grade on the U.S. history exam is lower than on the geography exam. This *suggests* that the U.S. history assessment was harder *and/or* that students performed less well, on average, than was the case for the Geography NAEP. For the Geography NAEP, students at the mean for Grade 4 got 42% of the total possible points on the items in the assessment, students at the mean for Grade 8 got 46% of the total possible points, and for Grade 12 students it was 48%. For the U.S. History NAEP, students at the mean for the Grade 4 assessment got 33% of the total possible points over all the items, students at the mean for Grade 8 got 38%, and students at the mean for Grade 12 got 35% of the total possible points on the items in the assessment for that grade level.

A comparison of the percent of total possible points at each cutscore to the percent of total possible points at the mean gives a sense of the *relative* distribution of student performance at each grade level. Clearly, the proportion of students whose scores correspond to approximately 75% of the total possible points on the Grade 12 U.S. history assessment is likely to be less than the proportion of students whose scores correspond to approximately 75% of the total possible points on the Grade 12 geography assessment, for example.

The cutscores at each grade remained relatively stable across the three rounds of ratings, while the variability (reported as the standard deviation) among raters decreased. One generally expects the panelists' understanding of both the rating task and of the achievement levels descriptions to increase throughout the process. One indicator that this was occurring is the decrease in variability among panelists' ratings.

Additional Analyses

The preliminary analyses have been focused on issues of interest or concern that emerged in relation to the 1990 and 1992 ALS processes.

Item Type

Significant differences in cutscores for dichotomous (multiple choice) items and polytomous (constructed response/essay) items were found. In all cases, the cutscores for polytomous items are higher than for dichotomous items. With the exception of Grade 4 Basic Geography, there is a statistically significant difference between polytomous and dichotomous item ratings.

Panelist Type

No significant differences were found among the cutscores set by teachers *versus* nonteachers *versus* general public panelists. This finding holds for both the geography and U.S. history ALS panelists.

While there is **no** significant difference in the levels set by panelists of the different types, teachers *generally* set the lowest levels for geography and non-teachers the highest. For U.S. history, the non-teachers again *generally* set the highest cutscores, but there is less consistency regarding the lowest levels. In any case, the differences are *not significant*.

Achievement Levels Descriptions

Panelists were trained in the framework and meaning of the preliminary achievement levels descriptions during the first two-three days of the ALS process. In addition to this thorough study of the preliminary descriptions, panelists were given opportunities to revisit the descriptions and make modifications before each round of ratings. After each session of review and proposed modification, the content staff met to discuss the suggestions and to determine jointly whether the suggestions seemed in keeping with the Framework, and whether the changes represented significant substantive changes. All changes for all grades were reviewed by all content staff. If the content staff agreed that the changes were consistent with the Framework, the changes were incorporated into the achievement levels descriptions and then taken back to the panelists for final agreement before each round of ratings. Panelists were advised that once the third (final) round of item ratings had begun, no further changes in achievement levels descriptions could be made.

Very few changes were recommended. In geography, only a few words changed at each the three grade levels. The only substantive change was at Grade 8 Basic with the addition of "and their physical features."

More changes were made in U.S. history than geography, but still the changes were few. The preliminary descriptions in U.S. history seemed less carefully worded than for geography. The content resource staff for history were clearly less committed to the wording of the preliminary descriptions than were the geography content staff. The changes appear to have improved the meaning of the descriptions and to have brought greater structural integrity across the grades.

The complete text of descriptions (along with selected items to illustrate each) for each grade level is included for geography and for U.S. history as Figures 2 and 3, respectively. Panelists have recommended that the achievement levels descriptions be made more "user friendly" in format. They recommend that some highlighted or "bulleted" text be extracted from the achievement levels descriptions to make the key aspects of the descriptions clear to readers.

Exemplar Items

Exemplar items were selected from the entire item pool at each grade level. Blocks had been identified for release in each content area. Project staff complied with the request of NAGB staff to have the exemplar items selected from all item blocks—not limited to those specific blocks "marked" for release by the National Center for Education Statistics and the Educational Testing Service, the NAEP contractor.

Exemplar items reviewed by panelists were pre-selected using statistical criteria. The statistical criteria were recommended by the Technical Advisory Committee on Standard Setting (TACSS). The procedure was first implemented during the U.S. history pilot study in August, 1994. Only items having an average probability of correct response of at least 50% for scores within the achievement level range were included in the list of items submitted to panelists for their consideration. Items were listed at the lowest achievement level for which the statistical criterion was met.

Constructed response items, scored for partial credit with 1-3 or 1-4 score points, were treated as separate items. For example, a score of "2" on one item could have met the statistical criterion for classification at the Basic level, whereas a score of "3" would have met the criterion from classification at the Proficient or Advanced level. Each "correct"[8] score point (2, 3, or 4) was evaluated to determine whether the average probability of that score point was ≥50% within the numerical range of each achievement level.[9]

From this list of items meeting the statistical criterion, panelists were asked to review items and to *veto* or eliminate those that would not serve to illustrate the knowledge and skills associated with performance for each achievement level. Because there were many items to review within each grade level, table groups of 5 panelists each were assigned one or two blocks of items to review. The table groups then made recommendations to the entire grade group. Any panelist within the grade group could recommend additional items to eliminate or to "reinstate." Items could be recommended as exemplars for a higher level, but not for a lower level. This means that an item that met the statistical criterion for consideration at the Basic level, for example, could be recommended as a Proficient level exemplar item—or even as an Advanced level exemplar item. An item that met the statistical criteria for consideration at the Proficient level could not, however, be recommended as a Basic level exemplar item.

Tables 4a and 4b present the counts of items on the pre-selection list meeting the statistical criteria and the counts of items recommended by panelists at each grade level. A few of the items recommended as the best of those marked for release are presented in Figures 2 and 3 along with the achievement level descriptions for which the items serve as "exemplars."

Panelist Evaluations

Understanding of Achievement Levels Descriptions

Analyses of the process evaluation questionnaires provide more information regarding the panelists' confidence in and satisfaction with the process. Responses to questionnaires were reviewed on site, and no process problems were observed. Panelists generally seemed to gain greater confidence in and satisfaction throughout the process. The data reported in Figures 4a and 4b show that their understanding of the definition of the achievement levels increased for ratings across rounds.

---

[8] Constructed response items that are "inappropriate" are coded "1." Those responses were not evaluated for consideration as exemplar items.

[9] A programming error led to the omission of all "items" with the highest scores from the list of items from which panelists made their recommendations. This error was not detected until March 1994. Panels representing one-third of the panelists at each grade level and for each subject were convened May 21-22 in Iowa City. After a brief reorientation to the achievement levels descriptions and Frameworks, these panelists reviewed all items in the item pool for their grade, reviewed the entire set of items that had been recommended during the achievement levels-setting meetings in November and December, and then reviewed those items recommended for the "released" blocks. They made recommendations regarding the items that had been previously omitted.

Conception of Borderline Performance

The percentages of panelists who reported that their conception of borderline performance was more than *moderately well formed* were generally lower than the percentages who reported that their understanding was more than *somewhat clear*. Figures 5a and 5b show that the conception of borderline performance gained clarity across rounds.

General public panelists were less likely than other panelists to have gained a well formed conception of borderline achievement prior to the third round of ratings. The general public panelists in geography were less likely to report a *very well formed* conception of borderline performance for each achievement level for ratings in Rounds 1 and 2. By Round 3, their responses indicated that their conceptions of borderline performance were about as well formed as those of other panelists.

Similarly, the U.S. history general public panelists responded less often that their conceptions of borderline performance were *very well formed* until after the second round of ratings. By the second round, their reports were similar to those of other panelists regarding their conception of borderline performance.

Rating Methods

The *conceptual clarity* and *ease of applying* increased for both the modified Angoff method for rating multiple choice items and the mean estimation method for rating constructed response items. While panelists reported increases in both dimensions for rating items of both formats, the *conceptual clarity* and *ease of applying* the modified Angoff method slightly exceeded the mean estimation method throughout the process. See Figures 6a and 6b for responses regarding the conceptual clarity of the methods, and Figures 7a and 7b for the responses regarding the ease of applying the methods for item ratings.

Differences in Performance Estimates by Item Type

Differences in achievement levels that would be set for ratings on multiple choice items only *versus* constructed response items only were consistently observed for rating data collected during the 1992 ALS process for mathematics and reading. And, the pilot study results for geography and U S. history revealed the same pattern. Given those observed differences, questions were constructed to collect information from panelists regarding their opinions related to student performance on multiple choice and constructed response items.

Responses indicated that panelists were more likely to attribute differences in performance ratings on the two types of items to differences in rating methodology in the beginning round of ratings. By the final round of ratings, however, they were more likely to *disagree* with this cause for differences in performance ratings than with *differences in student behavior and performance on the items* as the cause. This suggests that as panelists gained understanding, clarity, confidence, and so forth in the rating process, they became more convinced that such differences must be due to factors *other than* the rating methods.

Across each round of ratings, panelists were highly likely to agree that *constructed-response items assess dimensions of knowledge and skills that are significantly different from those*

assessed by multiple-choice items. Across the rounds of ratings, the responses tended to shift toward the *totally agree* response.

In the final "overall" process evaluation, panelists indicated that the instructions were clear and their level of understanding was adequate. The average response on a 5-point scale for panelists in both subject ALS processes was at least 4. The amount of time given to complete the tasks was about right. At least three-quarters of the panelists in both geography and U.S. history also indicated that it was more than *somewhat likely* that they were using the same concept of borderline performance as other panelists for all levels of achievement. Most panelists were more than *somewhat confident* in their ratings, and they were equally confident that the process was effective. About half of the panelists (51% for geography and 48% for U.S. history) reported that during discussions they felt a need to defend the ratings they had given *to some extent*. Few felt coerced to modify ratings from previous rounds. Most felt that the process provided them an opportunity to use their best judgment in rating items to set achievement levels. Further, most felt that the process produced *defensible* and *reasonable* achievement levels and that the panelists were *representative* and *credible*. Responses to items related to the above-mentioned factors averaged around 4 points or higher on a 5-point scale.

Consequences Data

Following the final round of ratings, all feedback data were updated and reported to panelists. In addition, data on the *consequences* of their ratings were provided. The *consequences data* were estimates of the percentages of students who would score at or above each achievement level cutscore at each grade. Panelists at each grade level were asked to comment on the data specific for their grade.

No such information had been collected from panelists during the 1992 ALS process. When the data became available to relate rating data to student performance on the NAEP scale, there was no way of knowing how the panelists would have reacted to the results or consequences of their ratings. These data were presented to panelists for their evaluation to inform decisions that would ultimately need to be made regarding the achievement levels. The data presented to panelists were not exact, and panelists were informed of this fact. The data were estimates based on a normal distribution rather than the actual distribution of scores on the NAEP scale.[10]

In general, the consequences data were consistent with panelists' expectations. Responses indicated that 89% of the geography panelists and 72% of the U.S. history panelists felt that the data reflected their expectations about the proportions of students at their grade level whose NAEP score would be at or above the cutscore for each achievement level. Five of the nine geography panelists who had not expected the results would change one or more of the achievement levels. In general, those panelists would opt to raise the cutscore in order to reduce the percentage of students scoring at or above a level. The Basic level was where most (4) felt the percentage should be smaller, and 2 of the 4 responses were for Grade 4 Basic.

[10] The NAEP scale distribution data were not available until April 1995. These data can only be reported by the Commissioner of Education Statistics, or some person authorized by the Commissioner.

For U.S. history, most panelists (13 = 57% of the 23 who did not expect those results) did *not* feel that the achievement levels should be changed in order to adjust the percentages of students scoring above the achievement levels. Of those 13 panelists who would make some changes, the general recommendation would be to lower the cutpoint to increase the percentage of students scoring at or above the level. This recommendation was made for all levels at Grades 4 and 8. Only one person recommended one change in the 12th grade levels, and that change would be to *raise* the cutscore in order to reduce the percentage of students scoring at or above the Basic level.

Conclusions

The achievement levels-setting process implemented for both the Geography NAEP and the U.S. History NAEP went very smoothly and according to plan. The success of the meetings was in large part attributable to the fact that complete pilot studies had been conducted for each subject.

The change that on-site staff recommended following the ALS meetings was that panelists not be asked to read student papers for all constructed response items in their rating pool. Staff felt the primary purpose of that task could be met by having panelists review far fewer papers leaving time to discuss the papers more fully with other panelists. Otherwise, no changes in the process appeared to be warranted.

Validation Research Studies

Similarity Classification Study

Overview

The Similarity Classification Study (SCS) is a research study involving teachers and their students. Eighth grade teachers[11] who served as panelists in either the pilot study or the achievement levels-setting (ALS) meetings for U.S. history and geography were asked to participate in the study. Students in up to two classes of each of these teachers were to be involved in a special assessment. The purpose was to determine whether and how well teachers who had participated in the setting of achievement levels could apply those descriptions of achievement to the abilities and likely performances of their own students. Having teachers who were very familiar with the assessment framework, the assessment item pool, the achievement levels descriptions, **and** students taking a course in the subject provided a *best case test*. These teachers would be asked to estimate the similarity of their students' knowledge and skills in the subject (geography or U.S. history) to the descriptions of what students should know and be able to do at each achievement level.

Teachers sent the roster of students in their classes slated to participate in the study. The total number of students for the geography teachers was 965, and 853 students were in the designated classes of the U.S. history teachers.

Students were administered a special form of the NAEP for the subject. This special, longer form was developed for this research study in order to get a reliable and valid measure of student ability over the entire content of the NAEP, as described by the Framework. The test form administered was twice the length of the NAEP: 100 minutes of cognitive assessment. Student background data and other student questionnaire data (courses in the subject, motivation on the exam, and so forth) were collected as well.

Teachers

Teachers were convened for their input March 4-6 in Kansas City, Missouri. Nineteen teachers of geography and 16 teachers of U.S. history participated in the study[12]. The teachers were provided a brief re-orientation to the framework and achievement levels descriptions for their subject. The content facilitators who had taken the lead in presenting the frameworks during the pilot and levels-setting meetings were present to provide this instruction.

Training for the task included, but was not limited to, a review of all items and the scoring rubrics in the eighth grade assessment. Teachers were told that the assessment to be administered to their students would be twice the length of a National Assessment, i.e., 100 minutes; but they were not told the number of item blocks nor the items to be administered.

---

[11] Seven geography teachers taught at grade 7 and two at grade 9. A total of 355 students were in the 7th grade and 126 in the 9th.

[12] The total number of geography teachers who served as eighth grade panelists in both the pilot study and the ALS was 31. The total number of U.S. history teachers who served as eighth grade panelists in the pilot study and the ALS was 26.

Teachers were instructed that they would estimate the similarity between each student and the three achievement levels. They were told that there would be two such similarity classifications, and that specific training and instructions would be provided for each session. During the first similarities classification session, they estimated the achievement of their students with respect to the framework and achievement levels descriptions of what students should know and be able to do. During the second similarities classification session, they estimated the achievement of their students with respect to the framework and achievement levels descriptions (as in the first classification), and they were instructed to take into account what they knew about the students and about the NAEP to be administered in making estimates about the students' performance.

Spaces were provided for describing factors taken into account in the second classification[13]. Most frequently mentioned factors were related to student behavior. **The** most frequently mentioned factor was the positive motivation of the student: *this student tries hard; will take the test seriously and try; very focused.* The *high academic skills and abilities* of the students were mentioned as often as *unpredictable; immature; can do well if tries.* Lack of motivation was the third most frequently mentioned factor: *does not try hard; will not / may not take this test seriously* were comments offered to describe student behavior influencing the teacher's estimate.

Students

The students were administered an "expanded" version of the NAEP for the subject. Students in all 35 schools were assessed, and scoring of the data was completed, as scheduled, by the end of April. Of the 1803 students to be assessed, 1649 were actually assessed.[14] A total of 154 students were absent and 10 students were "excluded," according to the usual NAEP administration procedures.

The expanded version of the NAEP included four blocks of items that were recommended by ETS, the NAEP contractor for test development, scaling, and most aspects of the assessment. The item blocks were selected to provide "equivalent" coverage of the NAEP item pool content and to be as similar as possible to the characteristics of the complete item pool, while minimizing the amount of time needed to test the students.

Additional Information

Teachers were asked to complete the teacher questionnaire and the school principal (or other appropriate administrative official) was asked to complete the school questionnaire. The data reported in these questionnaires will be helpful in comparing the relationships between student performances in this special study and that for the National Assessment.

---

[13] Six student forms were identified, and the teachers were instructed to describe the factors taken into account when estimating the achievement level to represent the performance for each of the six students. Two students classified at each achievement level (Basic, Proficient, and Advanced) were identified for these comments.

[14] Some of the students assessed were not included on the roster of students submitted by the teacher. Thus, the final count of students for which complete data, i.e., both teacher estimations of performance and actual performance data, will be somewhat less than this number of students assessed.

Findings

Only preliminary findings are available at this time. Comparisons of the two classifications by teachers revealed that teachers were more likely to estimate performance on the first classification as higher than on the second. That suggests that teachers believed that the ability of their students was higher than the students' performance on the special form of the NAEP was likely to reveal. Table 5 shows that only 11 percent of the second classifications by the geography teachers and 11 percent of the second classifications by the U.S. history teachers were higher than the first classifications. Geography teachers were most likely to make no change in the two classifications, while U.S. history teachers were most likely to lower the level for the second classification taking account of student behavior and the information available to them about the special NAEP.

No student performance data are available at this time. The expectation is that teachers estimated student ability and student performance to be at a higher level of achievement than test performance would indicate. The computations of individual student scores and comparisons of those scores to the teacher's classifications has not yet been completed.

Booklet Classification Study

Overview

The Booklet Classification Study (BCS) is conceptually similar to the Similarity Classification Study (SCS) described above. Rather than estimating the similarities of "real" students to the descriptions of what students should know and be able to do at each achievement level, however, the estimates were made of the similarity of the performance of a student on a NAEP test booklet to the achievement levels descriptions. Panelists did not know the students who took the test; they could only read the responses the students gave to the questions administered on the particular booklets.

The task was for panelists to classify the booklets into one of the achievement levels, based on an evaluation of the performance of that booklet relative to the descriptions of achievement at each of the three levels—plus the implicit "Below Basic" level.

Panels

The Booklet Classification Study was conducted for both geography and U.S. history in St. Louis, April 8-11. Ten panelists were invited to participate at each of the three grade levels for each subject. Approximately six teachers, two nonteacher panelists, and two general public members were represented on each grade-level panel.

Panelists for this study were selected on the same basis as panelists for the pilot studies and achievement levels-setting (ALS) panels. Previously-selected panelists who had been unable to serve on either the pilot or ALS panel were again asked to participate in this study. Additional persons with the same qualifications as the ALS panelists were recruited to serve in order to fill the panels with ten members each.

The Process

The training provided to panelists was very similar to that provided to the ALS panelists. The first two days of orientation, training, and practice exercises were modified only slightly from the sessions for the ALS process.

Staff were the same as for the ALS process, as well. Two content experts who had served with the previous pilot and ALS panels worked with these panelists to help them understand the frameworks and to gain a confident understanding of the achievement levels descriptions. Process facilitators who served in both pilot studies and ALS meetings again led panelists through the training and the process.

The Task

The task was for panelists to classify 40 NAEP booklets into the three achievement levels—or the "Below Basic" level. The booklets had been carefully selected according to statistical criteria. Namely, booklets were selected for which at least 4 of the 5 plausible values were

within the achievement level range.[15] A special effort was made to reduce the number of different items with which panelists would need to become familiar in classifying booklets.

The criteria used for selecting blocks of items for this study were the same used for the similarity classification study. Namely, the item characteristics for the items in the blocks selected matched as closely as possible those for the complete NAEP item pool. The distribution of items with respect to content categories and item formats were also approximately equal. Given that, the minimum number of different booklet forms, i.e., combinations of two blocks of items containing the four blocks identified for the study, were selected for the study. Specific booklets meeting the performance criteria (4 of 5 plausible values at the level) were then identified from among the booklets of each form.

Panelists generally had no difficulty completing the task within the allocated time. All history panelists had completed the task within approximately five hours. The geography panelists did not work so rapidly, and the last panelist completed the tasks after over seven hours.

Findings

The analyses will be conducted to determine the correspondence between booklet classifications and student performances. There is no expectation that there will be a perfect correspondence between the two. As indicated in the proposal for the study, however, "evidence in support of the achievement levels would be a reasonable level of relationship between the classification based on the NAEP scale data and the classifications by panelists, and a low level of shift in the assignment to categories." A "low level of shift" means, for example, that the panelists classified the booklets within one achievement level of the NAEP scale classification.

Only very preliminary analyses of the data have been conducted at this time. Based on these analyses, it appears that panelists generally classified the booklets at a level lower than the empirical (plausible values) classification. This finding would indicate that the achievement levels were set too low.

Further studies of the actual booklets used in the BCS are underway. One hypothesis is that panelists were unable to treat items in the way they are treated in NAEP scaling. This would mean that treatment of items in the model used by panelists for classifying booklets was different than that in the Item Response Theory model used for scaling NAEP data. Data collected from panelists on the strategy they used for the task, observation on site, and results indicate that many—if not most—panelists used a non-compensatory model and that they did not treat "not reached" items as if they were not administered. These considerations

[15] The instructions from our Technical Advisors were to select booklets with 3 of the 5 plausible values within the achievement level range **and** the average of plausible values at least as great as that for the level. If those criteria could not be reached for the targeted number of booklets per level, then fewer booklets were to be presented to the panelists. The targeted distribution of booklets was 7 at Below Basic, 13 at Basic, 13 at Proficient, and 7 at Advanced. The target of 7 booklets at the Advanced level, given the various criteria, was not met for all grades. The "rule" was that no fewer than 5 booklets could be used for the Below Basic or Advanced levels. Additional forms were to be added to assure that a minimum of 5 booklets were included for each level. Thus, for some grades, fewer than 40 booklets were presented to panelists for classification.

would lead panelists to classify performances represented in booklets as lower than the IRT scaling would be likely to do. Until further research can be completed, however, this is only speculation.

Table 1a

Geography ALS Nominee Pool

	TR	NT	GP	Total
Grade 4	59	8	12	79
Grade 8	50	6	12	68
Grade 12	49	9	16	74
Total	158 (72%)	23 (10%)	40 (18%)	221

	Male	Female	Total
Grade 4	13	66	79
Grade 8	34	34	68
Grade 12	37	37	74
Total	84 (38%)	137 (62%)	221

	White	Minority	Total
Grade 4	74	5	79
Grade 8	58	10	68
Grade 12	64	10	74
Total	196 (89%)	25 (11%)	221

	Ratings		
	1's	2's	3's
Grade 4	26	18	35
Grade 8	15	27	26
Grade 12	24	19	31
Total	65 (29%)	64 (29%)	92 (42%)

Table 1b

U.S. History ALS Nominee Pool

	TR	NT	GP	Total
Grade 4	29	12	20	61
Grade 8	37	17	20	74
Grade 12	43	15	26	84
Total	109 (50%)	44 (20%)	66 (30%)	219

	Male	Female	Total
Grade 4	15	46	61
Grade 8	35	39	74
Grade 12	46	38	84
Total	96 (44%)	123 (56%)	219

	White	Minority	Total
Grade 4	42	19	61
Grade 8	60	14	74
Grade 12	74	10	84
Total	176 (80%)	33 (20%)	219

	Ratings		
	1's	2's	3's
Grade 4	20	20	19
Grade 8	26	26	22
Grade 12	29	36	19
Total	75 (34%)	82 (37%)	60 (27%)

Table 2a

Geography Panelists

	TR	NT	GP	Total
Grade 4	17	5	8	30
Grade 8	18	3	7	28
Grade 12	18	4	9	31
Total	53 (60%)	12 (14%)	24 (27%)	89

	Male	Female
Grade 4	6	24
Grade 8	11	17
Grade 12	13	18
Total	30 (34%)	59 (66%)

	White	Minority
Grade 4	27	3
Grade 8	24	4
Grade 12	27	4
Total	78 (88%)	11 (12%)

Northeast	16	18%
Southeast	25	28%
Central	22	25%
West	26	29%

	Ratings		
	1's	2's	3's
Grade 4	2	5	23
Grade 8	2	6	20
Grade 12	2	8	21
Total	6 (7%)	19 (21%)	64 (72%)

Table 2b

U.S. History Panelists

	TR	NT	GP	Total
Grade 4	14	5	7	26
Grade 8	14	4	7	25
Grade 12	16	4	6	26
Total	44 (57%)	13 (17%)	20 (26%)	77

	Male	Female
Grade 4	9	17
Grade 8	6	19
Grade 12	13	13
Total	28 (36%)	49 (64%)

	White	Minority
Grade 4	19	7
Grade 8	20	5
Grade 12	23	7
Total	62 (81%)	15 (19%)

Northeast	16	21%
Southeast	21	27%
Central	21	27%
West	19	25%

	Ratings		
	1's	2's	3's
Grade 4	5	9	12
Grade 8	2	13	10
Grade 12	1	13	12
Total	8 (10%)	35 (45%)	34 (44%)

Table 3a

Geography ALS Cutscores[16] and % Correct Data[17]

	Basic	Proficient	Advanced
Grade 4			
Round 1	149.0 (9.9)[18]	166.8 (4.2)	178.9 (3.9)
Round 2	148.2 (7.9)	166.5 (4.3)	178.6 (4.0)
Round 3	148.3 (7.5)	166.5 (4.3)	178.6 (4.0)
% Correct	33.3%	61.1%	80.8%
Grade 8			
Round 1	144.7 (5.5)	162.5 (5.3)	174.4 (7.1)
Round 2	148.1 (5.5)	163.9 (4.6)	176.4 (5.7)
Round 3	148.6 (5.9)	164.4 (4.7)	177.3 (5.4)
% Correct	37.0%	60.7%	79.9%
Grade 12			
Round 1	145.8 (9.9)	164.4 (7.3)	176.8 (8.6)
Round 2	148.2 (5.7)	165.7 (6.1)	182.0 (5.8)
Round 3	147.7 (5.2)	165.5 (5.8)	181.9 (5.3)
% Correct	38.0%	63.6%	83.6%

[16] Cutscores are on the ACT NAEP-like scale.

[17] % correct data are estimates of the percentage of possible points required for a score at the lower borderline of each achievement level. Read: "A student would have to get at least 83.6% of the possible points on the items to score at the Advanced level in Grade 12."

[18] Numbers in parentheses are standard deviations for the cutscores.

Table 3b

U.S. History ALS Cutscores[19] and % Correct Data[20]

	Basic	Proficient	Advanced
Grade 4			
Round 1	157.1 (7.3)	170.1 (3.8)	180.6 (5.1)
Round 2	155.1 (4.6)	168.4 (3.4)	177.8 (4.9)
Round 3	155.5 (4.8)	168.8 (3.4)	178.3 (4.6)
% Correct	33.1%	56.7%	75.2%
Grade 8			
Round 1	155.2 (6.8)	171.5 (4.1)	185.1 (6.8)
Round 2	154.2 (5.1)	171.1 (3.9)	184.5 (5.7)
Round 3	154.3 (5.5)	171.2 (3.7)	184.6 (6.2)
% Correct	37.1%	61.6%	80.7%
Grade 12			
Round 1	159.3 (9 0)	170.7 (3.1)	182.5 (5.2)
Round 2	159.6 (6.2)	171.3 (2.5)	182.9 (4.4)
Round 3	159.6 (4.9)	171.6 (2.6)	183.6 (4.0)
% Correct	41.5%	63.2%	82.3%

[19] Cutscores are on the ACT NAEP-like scale.

[20] % correct data are estimates of the percentage of possible points required for a score at the lower borderline of each achievement level. Read: "A student would have to get at least 83.6% of the possible points on the items to score at the Advanced level in Grade 12."

Table 4a

Geography Exemplar Items from Released Blocks Only

	Statistically Selected	Panelists Recommended
Grade 4		
Basic	26	21
Proficient	16	14
Advanced	16	9
Grade 8		
Basic	31	4
Proficient	22	6
Advanced	9	2
Grade 12		
Basic	44	9
Proficient	13	4
Advanced	6	2
Grade 12 (w/Q3G3)		
Basic	56	19
Proficient	19	7
Advanced	9	4

Table 4b

U.S. History Exemplar Items from Released Blocks Only

	Statistically Selected	Panelists Recommended
Grade 4		
Basic	26	5
Proficient	25	20
Advanced	8	7
Grade 8		
Basic	32	18
Proficient	18	10
Advanced	11	7
Grade 12		
Basic	46	18
Proficient	21	16
Advanced	6	5

Table 5

Changes in Classifications Within and Across Levels: SCS

Subject	Change[21]	Frequency	Percent
Geography	Negative	301	31
	No Change	555	58
	Positive	101	11
U.S. History	Negative	419	50
	No Change	324	39
	Positive	94	11

[21] Change = (Classification 2 - Classification 1)

Figure 1a

Outstanding Credentials for Geography Panelists

Teachers

National Awards
1994 Teacher of the Year (State Farm Insurance Company)
National Park Service Outstanding Educator
President's Award for Science Teaching
Geography Education Award
Who' Who in Teaching
Unspecified (3)

State Awards
Teacher of the Year (5)
Governor's Awards/Citations for Teaching Excellence (2)
Montclair Award (CBS)
Other/unspecified (7)

Geography Alliance
23 Teachers
6 Nonteacher Educators

Nonteacher Educators

Local School/District Administrators (7)
State Administrators (1)
Post-Secondary Faculty (5)

General Public

Mayor
Owner map store
State Engineer
Owner, Trucking Business
Director of State Division of Archaeology
Peace Corp; Korean experience
Geography major and President of a broadcasting company
VISTA volunteer; Human Resources; well-traveled
District Manager, worldwide communications company
City Planner and Director of Community Development Programs
Retired Military; much travel
Research Analyst; well-traveled
Retired Program Manager for Lockheed; well-traveled; ties to education
Architectural Historian; travel; cultural geography interests and training
Dairy farmer with 5 children; geography is special interest
Military; world traveller
CEO small credit unit

Figure 1b

Outstanding Credentials for U.S. History Panelists

Teachers

National Awards
DAR History Teacher of the Year
Fullbright to Turkey
NEH Fellowships (3)
U.S. Inst. of Peace Grant
Certified grants writer (secured a $5 million grant)
Who's Who Among Teachers
Consulting work for Georgia State Department of Education
Woodrow Wilson Fellowship for American History at Princeton University
Unspecified

State Awards:
Teacher of the Year (4)
Citations for Teaching Excellence (2)
Al SS Teacher
Chairperson of Curriculum Assessment in AL
Governor's Award for Science
Other/unspecified (2)

Memberships:
NCSS (14)
State level CSS (8)
Other (5)

Nonteacher Educators

Local/District Administrators (6)
State Administrators (4)
Post-Secondary Faculty (4)

General Public

Author (2)
Park Ranger; New England & SW History
Pediatric Occupational Therapy
Owner, video production; western history
Museum Director
Oregon Trail history
Computer Network Coordinator; Constitutional History
Business Manager, Capital Electric Corp.; World Wars & Space Programs
Owner, mortgage company; plans on masters in secondary education
Museum Director; degree in history, UCal-San Diego
Bookroom clerk at junior high; helps students and school with book selections
Supervisor, Air Traffic Tower; Historical Society
Executive Director, Chamber of Commerce; education
Child Welfare Administrator; education; school board; tutoring
Museum Director; local history; historical reading
Executive Director, Chamber of Commerce; history interest

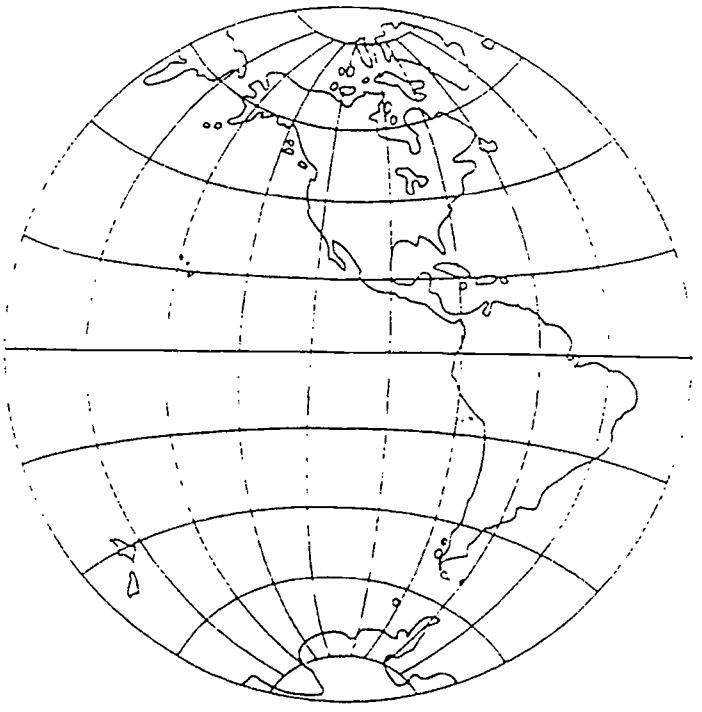
Church Secretary; school board; antiques; writes about antiques
Executive Director, Historical Society; Ph.D. Amer. Studies, Hist. & Envrn. History
Museum Curator; US History; French/Indian War; WW II; Scotish & Colonial history
Executive Director, Greater East St. Louis Chamber; world history
Owner, TMS Services; school board; education; history
Executive Director, Historical Society; college archeology/anthropology
Asst. Federal Public Defender; Political science major; student teaching in history

Memberships:
NCSS (8)
Other (1)

Figure 2

Geography Achievement Levels Descriptions and Exemplar Items

4th-Grade Basic—Students should be able to use words or diagrams to define basic geographic vocabulary; identify personal behaviors and perspectives related to the environment and describe some environmental and cultural issues in their community; use visual and technological tools to access information; identify major geographic features on maps and globes; be able to read and draw simple maps, map keys and legends; demonstrate how people depend upon, use, and adapt to the environment; and give examples of the movement of people, goods, services, and ideas from one place to another. In addition to demonstrating an understanding of how individuals are alike and different, they should demonstrate a knowledge of the ways people depend on each other.



11. On the map above, write the names of the North Pole, the South Pole, and the equator in the correct location.

KJ000664

Grade Level:	4
Achievement Level:	Basic
Block Number:	Q1G3
Item Number:	11 ≥3
Overall P-Value:	77
Avg Conditional P-Value:	59
Score Levels:	3
Content Area:	1

Rationale Text:
Scoring Guide

Scoring Rationale: Student locates and correctly labels the North Pole, South Pole, and the equator on a map showing the Western hemisphere.

3 - **Complete.** The response correctly labels all three features on the map.

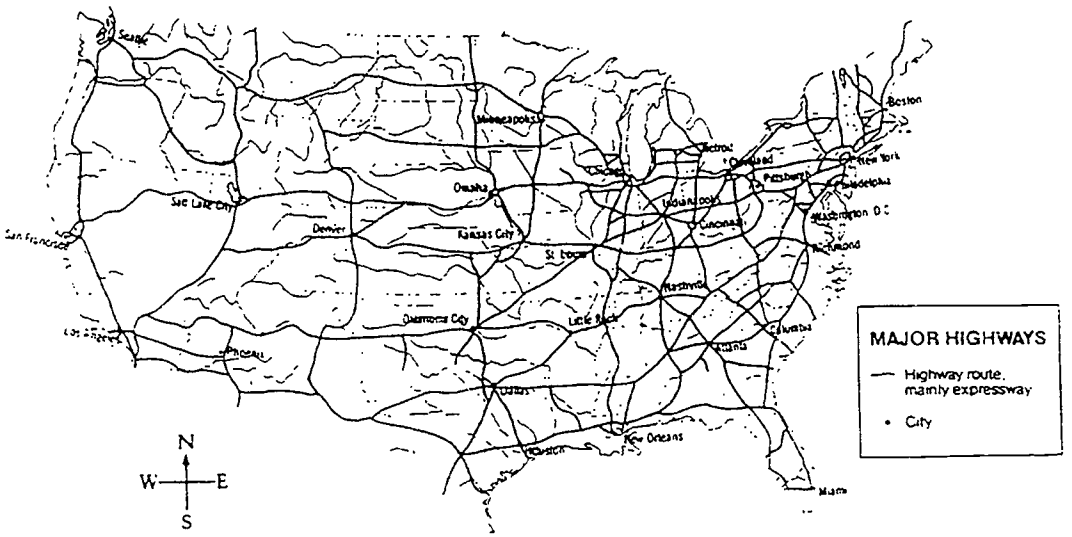
2 - **Partial.** The response correctly identifies the location of one or two features on the map.

1 - **Inappropriate.** The response does not correctly locate any of the three features.

Credited Responses:

- the North Pole at the top of the map
- the South Pole at the bottom of the map
- the Equator in the map's center.

Questions 12–13 are based on the highway map below. BO001961



12. The map shows that one part of the country has more major highways than the other part of the country. Why is this?

● There are more people and cities in the eastern part of the country.

Ⓑ It is easier to build highways in the eastern part of the country.

Ⓒ Cars are not an important form of transportation in the western part of the country.

Ⓓ States are larger in the western part of the country. BO001962

Grade Level:	4
Achievement Level:	Basic
Block Number:	Q1G6
Item Number:	12
Overall P-Value:	60
Avg Conditional P-Value:	59
Score Levels:	2
Content Area:	3

4th-Grade Proficient—Students should be able to use fundamental geographic knowledge and vocabulary to identify basic geographic patterns and processes; describe an environmental or cultural issue from more than one perspective; and read and interpret information from visual and technological tools such as photographs, maps and globes, aerial photography, and satellite images. They should be able to use number and letter grids to plot specific locations; understand relative location terms; and sketch simple maps and describe and/or draw landscapes they have observed or studied. Proficient students should be able to illustrate ways people depend upon, adapt to, and modify the environment; describe and/or illustrate geographic aspects of a region using fundamental geographic vocabulary and give reasons for current human migration; discuss the impact a location has upon cultural similarities and differences; and be able to demonstrate how a event in one location can have an impact upon another location.

Questions 6–7 refer to the following chart.

MAJOR EXPORTS OF THREE COUNTRIES

Country A	Country B	Country C
Oil	Cars	Computers
Natural Gas	Televisions	Airplanes
Coconuts	Cameras	Wheat

K1000759

6. The situation shown in the chart will probably lead to

● trade among all three countries

Ⓑ trade only between countries A and B

Ⓒ trade only between countries B and C

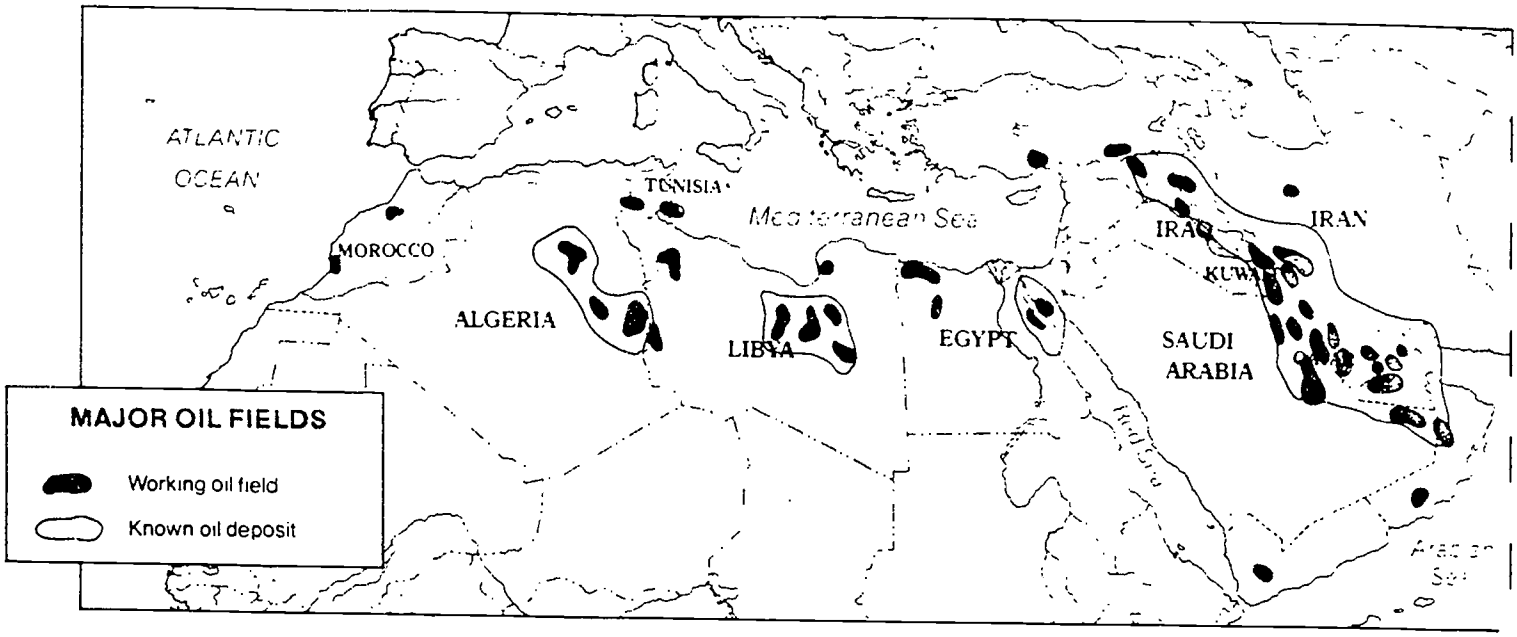
Ⓓ a decision by each country to produce all nine goods listed K1000760

Grade Level:	4
Achievement Level:	Proficient
Block Number:	Q1G6
Item Number:	6
Overall P-Value:	37
Avg Conditional P-Value:	67
Score Levels:	2
Content Area:	3

6. Look at the map on the top of page 65. What is a reason why the countries surrounding the Persian Gulf are important to many other countries in the world?

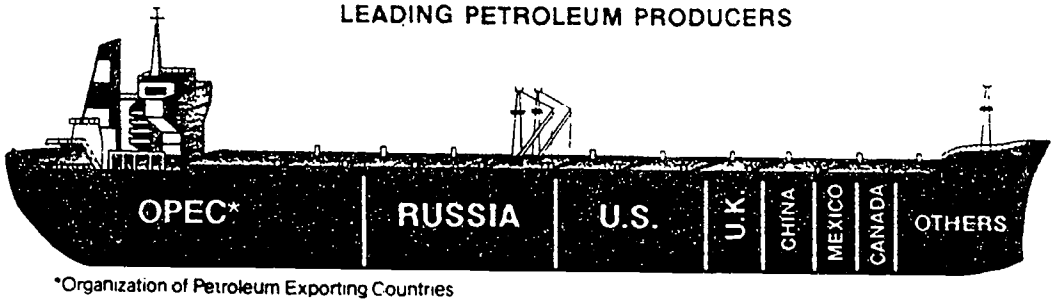
BO001942

Grade Level:	4
Achievement Level:	Proficient
Block Number:	Q12G7
Item Number:	6
Overall P-Value:	34
Avg Conditional P-Value:	63
Score Levels:	2
Content Area:	3



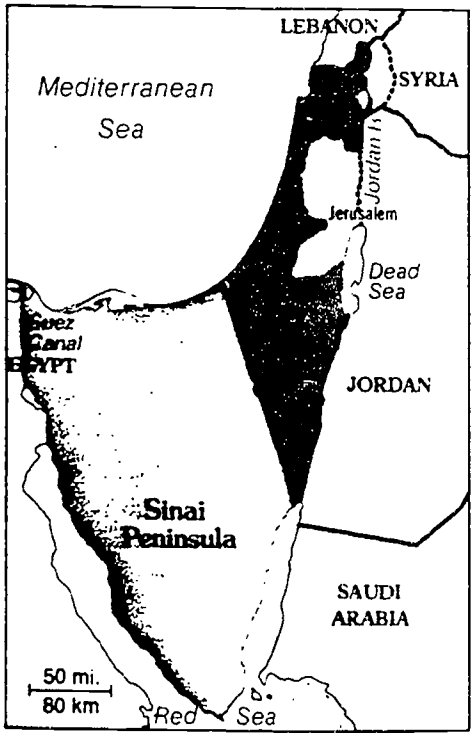
LEADING PETROLEUM PRODUCERS

Almost ⅓ of the world's petroleum is produced by OPEC.* OPEC members include the Persian Gulf countries of Iran, Iraq, Kuwait, Qatar, Saudi Arabia, and the United Arab Emirates. Other OPEC members are Algeria, Libya, Nigeria, Gabon, Indonesia, Ecuador, and Venezuela.

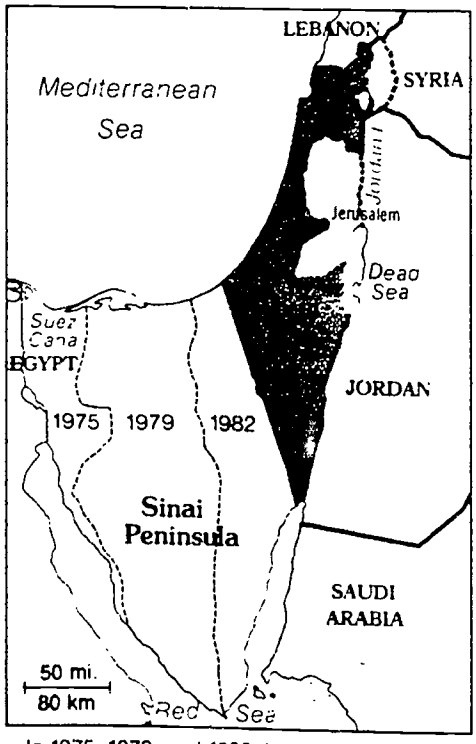


*Organization of Petroleum Exporting Countries

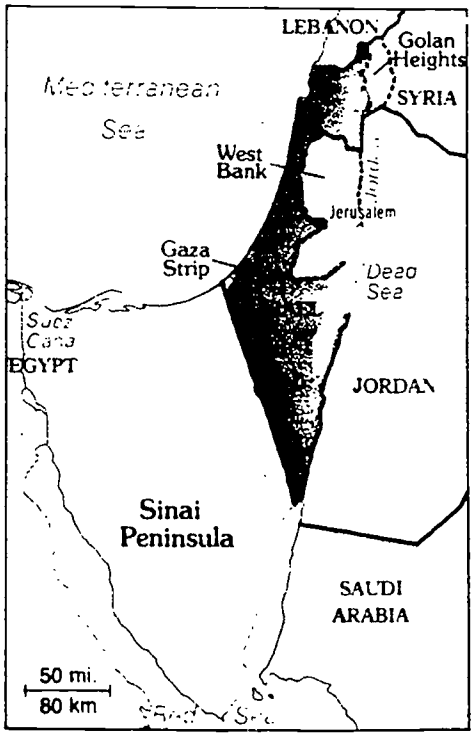
CHANGING BOUNDARIES



Israel occupied the area shown in dark green until 1967. After the Six Day War of that year, it also controlled the parts of Egypt, Jordan and Syria shown in light green.



In 1975, 1979, and 1982, Israel withdrew in stages from the Egyptian territory known as the Sinai Peninsula.



After withdrawing from the Sinai, Israel remained in control of the Gaza Strip, West Bank, and Golan Heights. In 1981 Israel claimed the Golan Heights as part of its state

Item Number: 6 Key: NONE Accession Number: BO001942

Classification Codes:
1/2 III B APPLY SA WRIT ATLAS

Look at the map on the top of page 65. What is a reason why the countries surrounding the Persian Gulf are important to many other countries in the world?

Rationale Text:

Scoring Guide

Scoring Rationale: Student demonstrates ability to interpret map of middle East.

3 - Complete. The response mentions the fact that the countries surrounding the Persian Gulf region are important because of the number of oilfields located there.

1 - Inappropriate. The response does not identify any reasons why the Persian Gulf is important to the rest of the world.

4th-Grade Advanced—Students should be able to use basic geographic knowledge and vocabulary to describe global patterns and processes; describe ways individuals can protect and enhance environmental quality; describe how modifications to the environment may have a variety of consequences; explain differing perspectives that apply to local environmental or cultural issues; and demonstrate an understanding of forces that result in migration, changing demographics, and boundary changes. They should be able to solve simple problems by applying information learned through working with visual and technological tools such as aerial and other photographs, maps and globes, atlases, news media, and computers. They should be able to construct models and sketch and label maps of their own state, the United States, and the world; use them to describe and compare differences, similarities, and patterns of change in landscapes; and be able to predict the impact a change in one location can have on another. They should be able to analyze the ways individuals and groups interact.

11. WAYS TO GET RID OF WASTE

- Dumping far out in the ocean
- Burning
- Recycling
- Burying in landfills

From the list above, select one method of getting rid of waste and identify one advantage and one disadvantage of this method. K1000826

Method of waste disposal: _____

Advantage: _____

_____

Disadvantage: _____

_____

Grade Level:	4
Achievement Level:	Advanced
Block Number:	Q1G6
Item Number:	11 ≥3
Overall P-Value:	33
Avg Conditional P-Value:	80
Score Levels:	3
Content Area:	2

Item Number: 11 Key: NONE Accession Number: KJ000826

Classification Codes:
1 II B APPLY SA WRIT NA

WAYS TO GET RID OF WASTE

- Dumping far out in the ocean
- Burning
- Recycling
- Burying in landfills

From the list above, select one method of getting rid of waste and identify one advantage and one disadvantage of this method.

Method of waste disposal: _____

Advantage: _____
_____

Disadvantage: _____
_____

Rationale Text:
Scoring Guide

Scoring Rationale: Student demonstrates ability to identify one advantage and one disadvantage of a selected waste disposal method.

3 - **Complete.** The response accurately describes an advantage and disadvantage of one method of waste disposal. Explanations should be both specific to that method and geographically logical. Correct answers may be drawn from the list given or include some other appropriate response.

2 - **Partial.** The response describes either an advantage or a disadvantage. Correct answers may be drawn from the list given or include some other appropriate response. If present, the other description is incorrect or trivial, as in, "dumping waste in oceans has no effect on us," or "it takes a long time to dump waste in oceans".

1 - **Inappropriate.** The response does not accurately describe an advantage or a disadvantage of a given method of waste disposal.

Credited Responses could include:

Ocean Dumping

Advantages

- Easy, low cost solution
- No costly land wasted for landfills
- No unsightly incinerators and landfills
- Requires no technology
- It can't be seen

Disadvantages

- Kills sea life (fish, birds, plants)
- Destroys fishing industry (especially shellfish)
- Ruins tourism (beaches closed for health and safety reasons)
- Water contamination
- Not all things are recyclable

Burning Waste

Advantages

- Reduces cost and waste of land used for landfills
- Energy created as a byproduct (steam and electricity)
- Recycling of useful products often accompanies burning
- can be economical option (in some cases, low waste transportation costs)

Disadvantages

- The production of ash and toxic gases
- Air pollution
- Costly to regulate and check safety of process
- Difficulty finding a location for incinerator
- Opposition to siting

Recycling

Advantages

- Produces less pollution than other methods
- Preserves natural resources (examples include wood, minerals, fuel)
- Reduces waste levels
- Reduce cost and waste of land used for landfills
- New uses for discarded items
- increases public awareness of waste and pollution
- encourages manufactures to produce

Disadvantages

- Achieving the motivation and cooperation required
- Requires public education and enforcement
- Need to find a market for recycled products
- Not all things are recyclable
- Needs initial equipment that may be expensive (economic factor)
- Other types of pollution may be produced during recycling

Burying Waste

Advantages

- Cheap, although only in the short term
- Simple
- Low technology
- In areas with low population density, can be popular with public

Disadvantages

- High costs of transporting waste to distant landfills
- Shortage of affordable and available land
- Contamination of water, soil, and air
- Smell
- Opposition to siting
- Not all products are biodegradable

7. Look at the population and landscape map of China on page 72 of the atlas.

What area of the country is densely populated (crowded)?

_____

_____

Give two reasons why this area of the country is densely populated. B0C01944

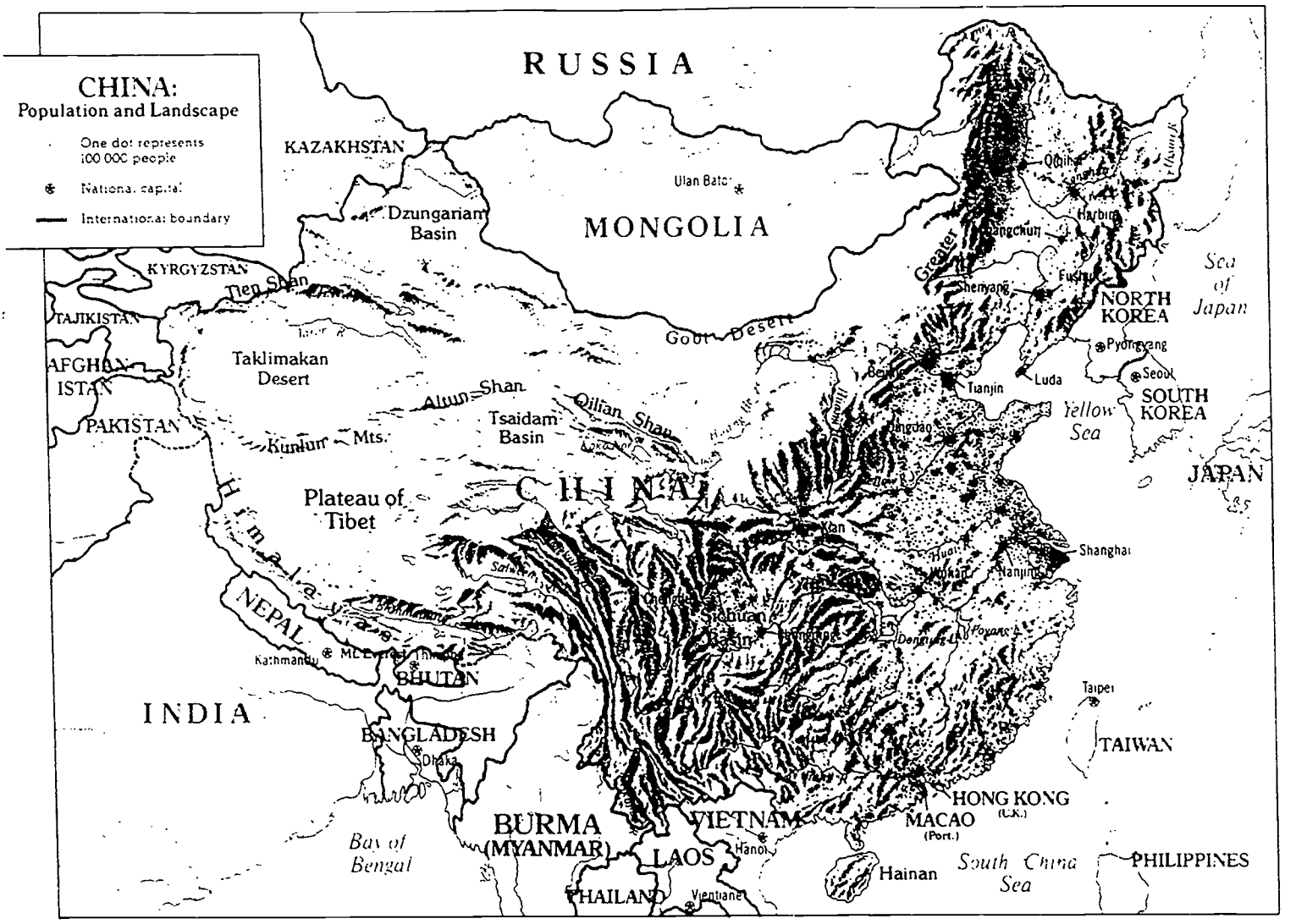
_____

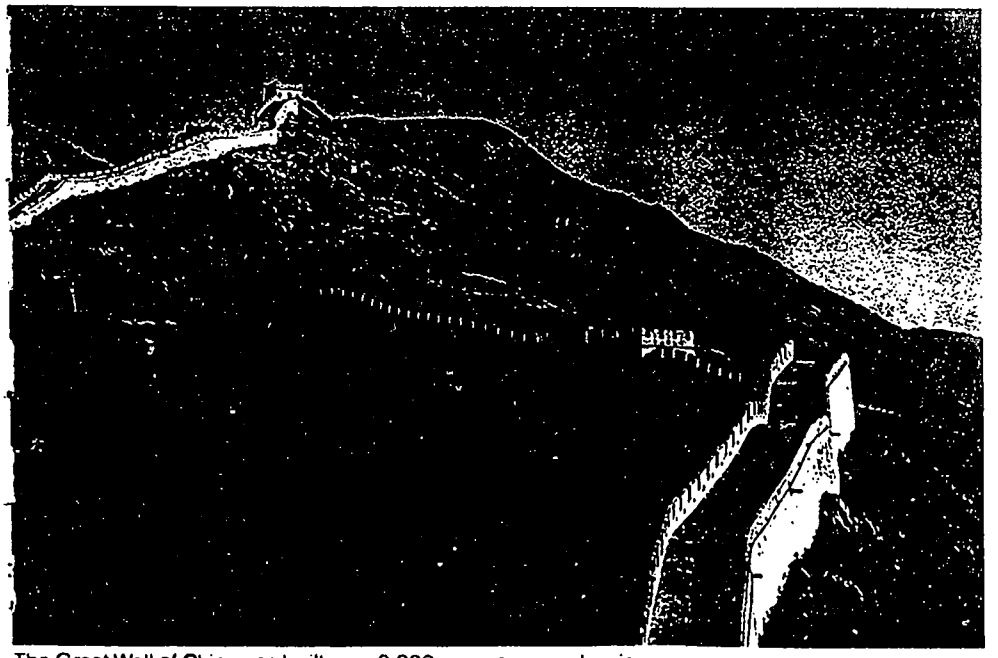
_____

_____

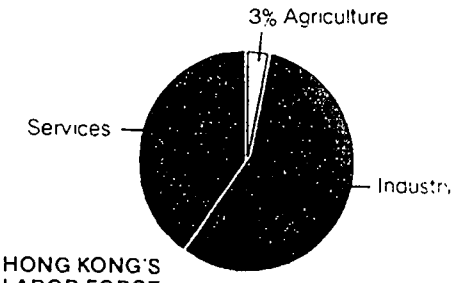
Grade Level:	4
Achievement Level:	Advanced
Block Number:	Q12G7
Item Number:	7 ≥3
Overall P-Value:	20
Avg Conditional P-Value:	79
Score Levels:	4
Content Area:	3

72 ASIA

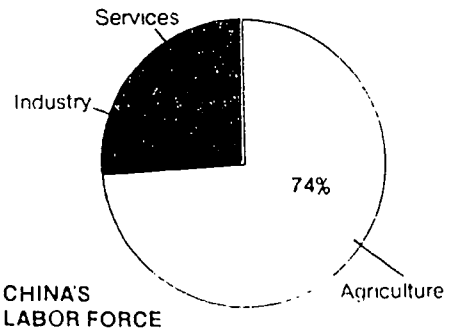




The Great Wall of China was built over 2,200 years ago as a barrier Mongol invaders. It remains the world's longest structure, g across 1,500 miles (2415 km) of northern China



HONG KONG'S LABOR FORCE



CHINA'S LABOR FORCE

Item Number: 7 Key: NONE Accession Number: B0001944

Classification Codes:
1/2 III A UNDER EA WRIT ATLAS

Look at the population and landscape map of China on page 72 of the atlas.

What area of the country is densely populated (crowded)?

Give two reasons why this area of the country is densely populated.

Rationale Text:
Scoring Guide

Scoring Rationale: Student displays an ability to interpret a population and landscape map of China by stating which area of the country is densely populated. The student also demonstrates an understanding of the causes of population density.

4 - **Complete.** The response correctly identifies an area of population concentration (the east coastal area or Sichuan Basin). It explains this by drawing on <u>two</u> of the explanations given in the list below or gives some other appropriate response.

3 - **Essential.** The response correctly identifies an area of population concentration (east coastal area or Sichuan Basin). It explains this by stating one of the explanations given in the list below or gives some other appropriate response.

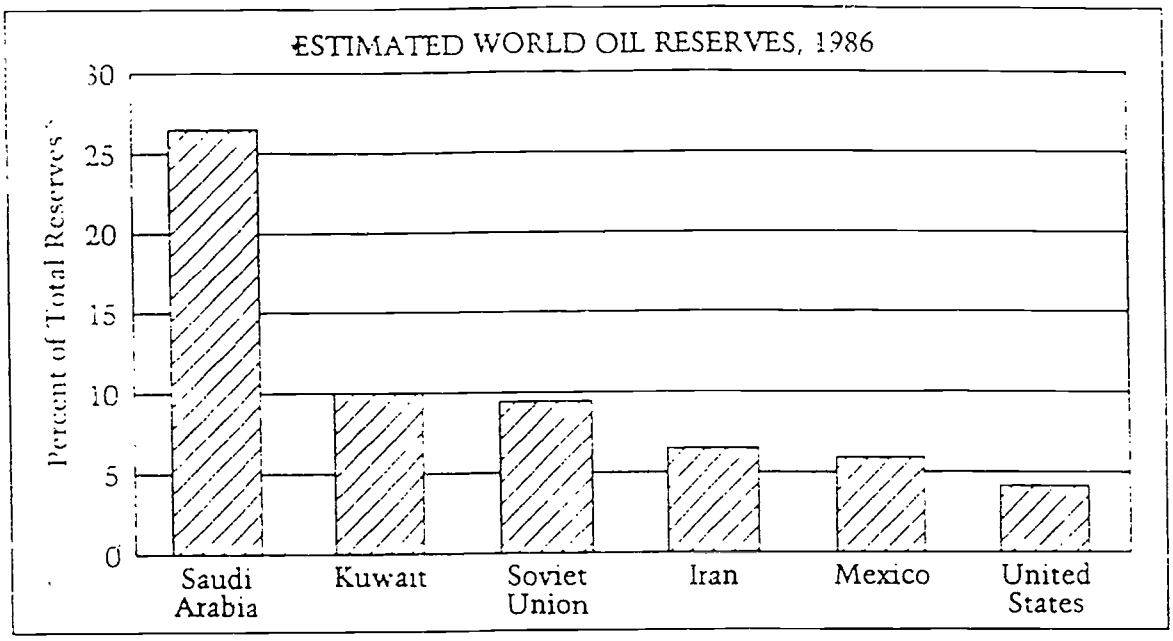
2 - **Partial.** The response correctly identifies an area of population concentration, but does not correctly explain why the population is concentrated in certain areas in terms of the explanations given or some other appropriate response.

1 - **Inappropriate.** The response does not correctly identify an area of population concentration as the east coastal area, or the Sichuan Basin, and does not explain the population pattern in terms of the explanations given below or some other appropriate response.

Credited Responses could include:

<u>Influences on Population Distribution</u>
- Coastal areas have flat land, suitable for building and agriculture
- The coastal area allows trade with other countries
- Oceans and rivers are important for settlement with a reason such as transportation, agriculture, household use, industry, fishing
- Much of China is mountainous or desert and is not suitable for settlement
- Government administrative center; capital (also major cities)
- the area has a lot of business and industry
- the area has a lot of agriculture

8th-Grade Basic—Students should possess fundamental knowledge and vocabulary of concepts relating to patterns, relationships, distance, direction, scale, boundary, site, and situation; solve fundamental locational questions using latitude and longitude; interpret simple map scales; identify continents *and their physical features*, oceans, and various countries and cities; respond accurately to descriptive questions using information obtained by use of visual and technological tools such as geographic models and/or translate that information into words; explain differences between maps and globes; and find a wide range of information using an atlas or almanac. Students should be able to recognize and illustrate the relationships that exist between humans and their environments, and provide evidence showing how physical habitat can influence human activity. They should be able to define a region and identify its distinguishing characteristics. Finally, they should be able to demonstrate how the interaction that takes place between and among regions is related to the movement of people, goods, services, and ideas.



KI000691

1. According to the graph above, the largest oil reserves in 1986 were in the

Ⓐ Gulf of Mexico

Ⓑ Caribbean region

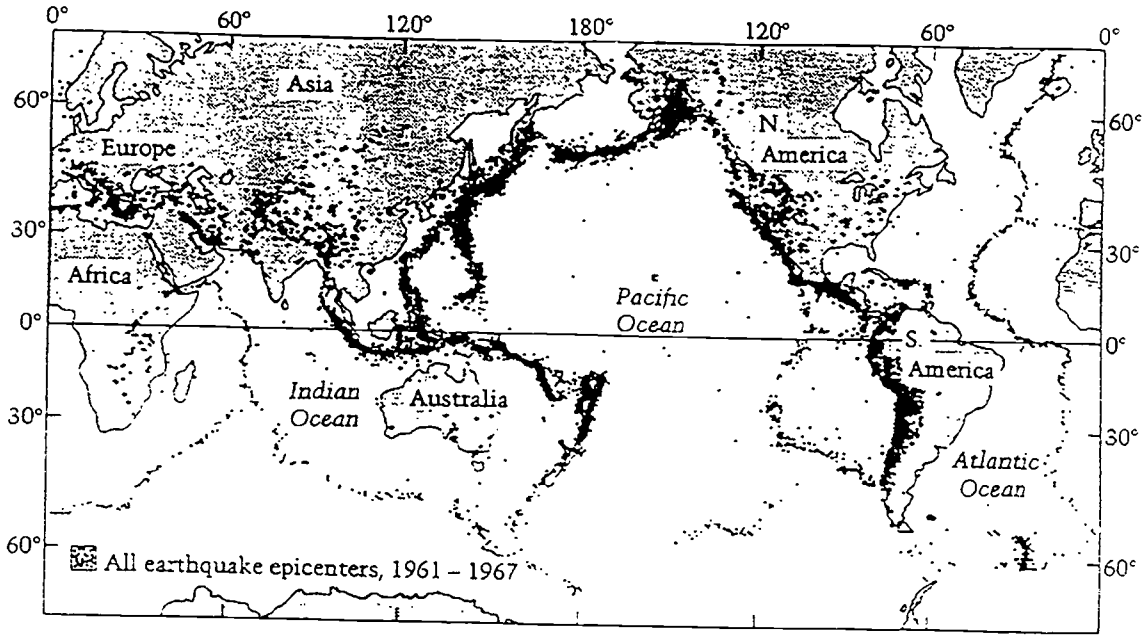
● Persian Gulf region

Ⓓ Gulf of Guinea

KI000692

Grade Level:	8
Achievement Level:	Basic
Block Number:	Q2G4
Item Number:	1
Overall P-Value:	85
Avg Conditional P-Value:	88
Score Levels:	2
Content Area:	1

Questions 3-4 are based on the map below showing the distribution of earthquake epicenters around the world between 1961 and 1967.



3. Between 1961 and 1967, the area that had the most earthquakes was the

Ⓐ Mediterranean basin

Ⓑ mid-Atlantic Ocean

Ⓒ Caribbean Sea

● Pacific Ocean rim

KJ00C79C

Grade Level:	8
Achievement Level:	Basic
Block Number:	Q23G6
Item Number:	3
Overall P-Value:	84
Avg Conditional P-Value:	89
Score Levels:	2
Content Area:	1

8th-Grade Proficient—Students should possess a fundamental geographic vocabulary; understand geography's analytical concepts; solve locational questions requiring integration of information from two or more sources, such as atlases or globes; compare information presented at different scales; identify a wide variety of physical and cultural features and describe regional patterns. Students should be able to respond accurately to interpretive questions using geography's visual and technological tools and translate that information into patterns; identify differences in map projections and select proper projections for various purposes; and develop a case study working with geography's analytical concepts. In addition, students should be able to describe the physical and cultural characteristics of places; explain how places change due to human activity; explain and illustrate how the concept of regions can be used as a strategy for organizing and understanding Earth's surface. Students should be able to analyze and interpret data bases and case studies as well as use information from maps to describe the role that regions play in influencing trade and migration patterns and cultural and political interaction.

8. The major areas of wheat production in the world are the central United States and Canada, Ukraine, south central Australia, and the pampas of Argentina. What is the characteristic shared by these areas that explains their role in wheat production?

Ⓐ All have rainy, damp climates.

Ⓑ All are near sea coasts.

● All are plains.

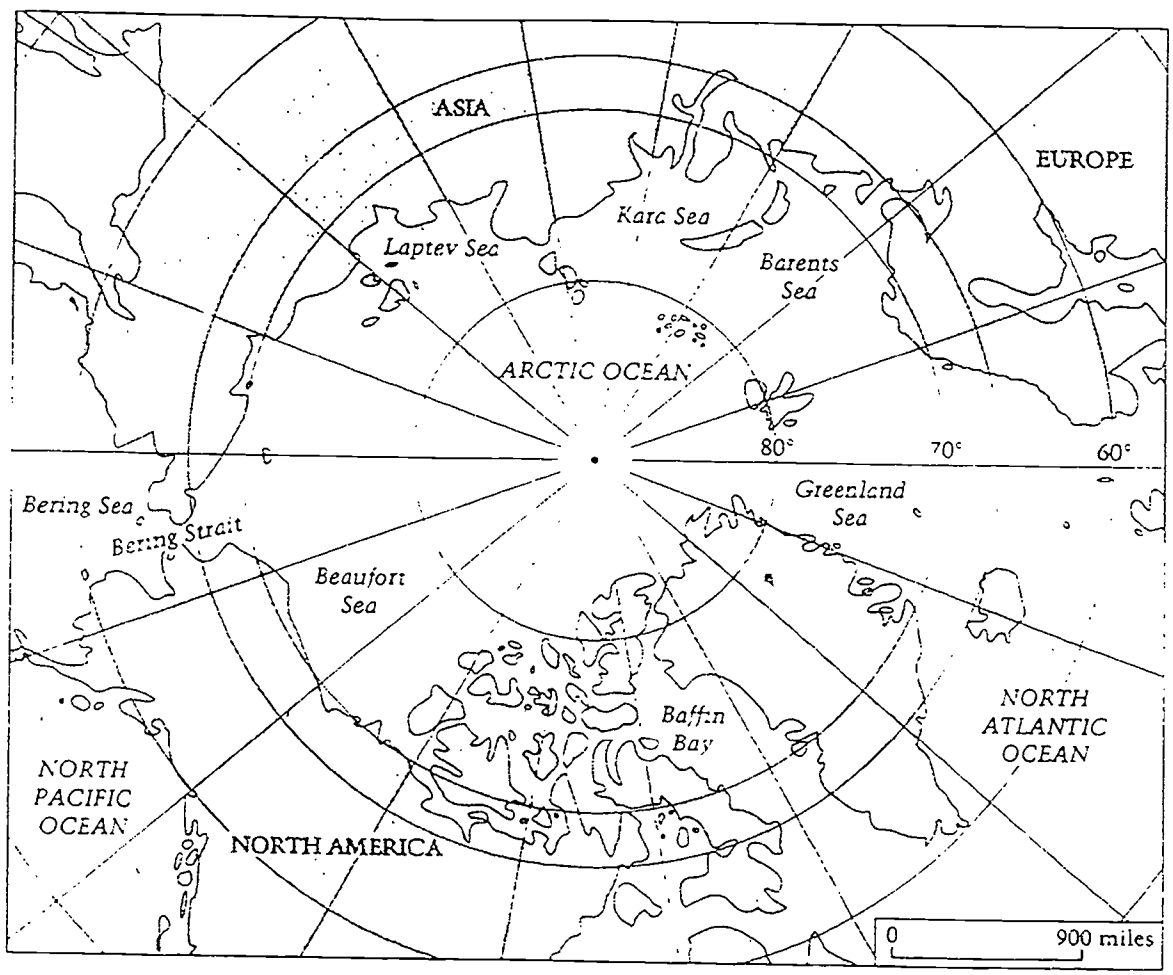
Ⓒ All are in highland regions.

K1000758

Grade Level:	8
Achievement Level:	Proficient
Block Number:	Q23G6
Item Number:	8
Overall P-Value:	52
Avg Conditional P-Value:	76
Score Levels:	2
Content Area:	1

Questions 16-17 refer to the map below.

SE0C068



16. This map would be most useful to a

Ⓐ pilot flying from Europe to South America

● pilot flying from Canada to Scandinavia

Ⓒ person sailing to Antarctica

Ⓓ person sailing in tropical seas

BO001953

Grade Level:	8
Achievement Level:	Proficient
Block Number:	Q2G4
Item Number:	16
Overall P-Value:	46
Avg Conditional P-Value:	80
Score Levels:	2
Content Area:	1

8th-Grade Advanced—Students should have a command of extensive geographic knowledge, analytical concepts, and vocabulary; be able to analyze spatial phenomena using a variety of sources with information presented at a variety of scales and show relationships between them; and use case studies for spatial analysis and to develop maps and other graphics. Students should be able to identify patterns of climate, vegetation, and population across Earth's surface and interpret relationships between and among these patterns, and use one category of a map or aerial photograph to predict other features of a place such as vegetation based on climate or population density based on topographic features. Students should also be able to relate the concept of region to specific places and explain how regions change over time due to a variety of factors. They should be able to profile a region of their own design using geographic concepts, tools, and skills.

3. What would a scientist probably study to predict where acid rain would fall?

Ⓐ The atomic structures of sulfur, nitrogen, and oxygen

Ⓑ Mass-transit systems that serve major cities

● Wind patterns that prevail over major manufacturing areas

Ⓓ The location of sewage-treatment plants

K1000617

Grade Level:	8
Achievement Level:	Advanced
Block Number:	Q2G4
Item Number:	3
Overall P-Value:	36
Avg Conditional P-Value:	84
Score Levels:	2
Content Area:	2

4. What are the conditions that make the tundra difficult for human settlement?

KIC008:

Grade Level:	8
Achievement Level:	Advanced
Block Number:	Q2G4
Item Number:	4
Overall P-Value:	75
Avg Conditional P-Value:	62
Score Levels:	3
Content Area:	1

Item Number: 4 Key: NONE Accession Number: KJ000824

Classification Codes:

2	II	B	UNDER	SA	NA	NA

What are the conditions that make the tundra difficult for human settlement?

Rationale Text:

Scoring Guide

Scoring Rationale: Student demonstrates knowledge of characteristics of tundra.

3 - Complete. The response identifies at least two characteristics of the tundra and may explain how these make it difficult for human settlement.

2 - Partial. The response identifies only one characteristic of the tundra and may explain how this makes life difficult for human settlement.

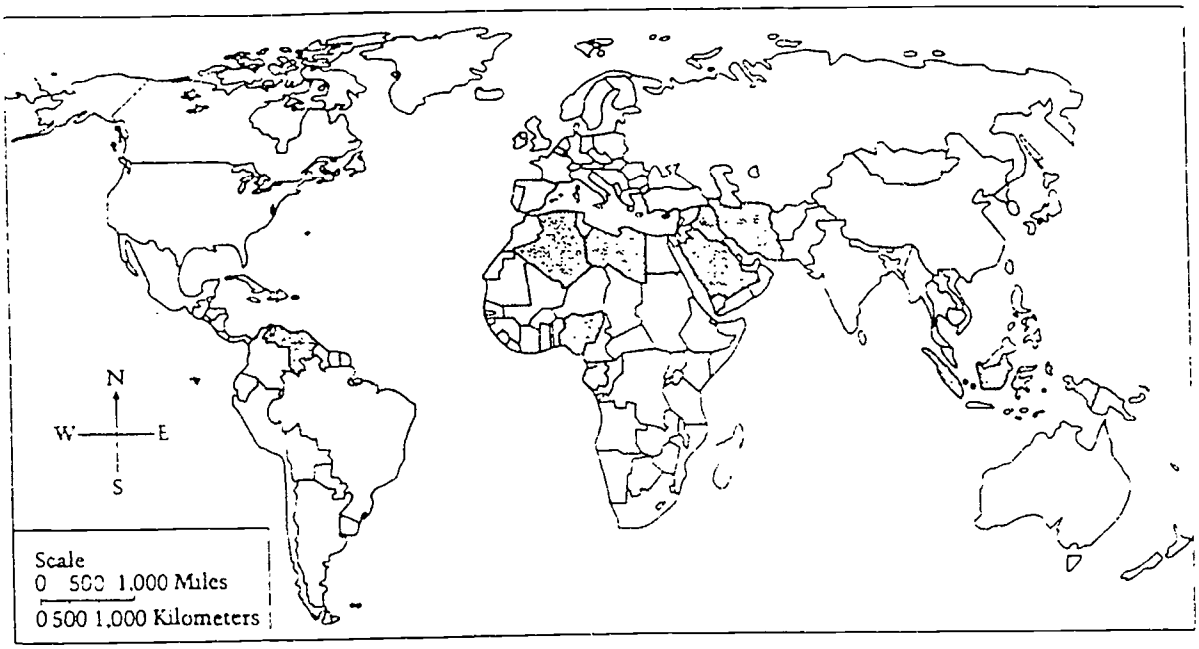
1 - Inappropriate. The response does not identify any characteristics of the tundra or explain how they make it difficult for human settlement. Or it provides answers that are inappropriate.

Credited Responses could include:

Tundra
- climate inhospitable
- very wet
- very cold
- permafrost
- frost almost all year
- lack of vegetation to support survival
- long hours of darkness
- lots of mosquitoes in summer
- lack of precipitation (generally low precipitation)
- icy gale-like winds

NB: Do not accept:
- land is dry
- weather is bad
- snow

12th-Grade Basic—Students should possess a knowledge of concepts and terms commonly used in physical and human geography as well as skills enabling them to employ applicable units of measurement and scale when solving simple locational problems using maps and globes. They should be able to read maps; provide examples of plains, plateaus, hills, and mountains; and locate continents, major bodies of water, and selected countries and cities. They should be able to interpret geographic data and use visual and technological tools such as charts, tables, cartograms, and graphs; know the nature of and be able to identify several basic types of map projections; understand the basic physical structure of the planet; explain and apply concepts such as continental drift and plate tectonics; and describe geography's analytical concepts using case studies. Students should have a comprehensive understanding of spatial relationships including the ability to recognize patterns that exist across Earth in terms of phenomena, including climate regions, time zones, population distributions, availability of resources, vegetation zones, and transportation and communication networks. They should be able to develop data bases about specific places and provide a simple analysis about their importance.



3. On the map above, the shaded countries represent the membership of the

   ● Organization of Petroleum Exporting Countries (OPEC)

   Ⓑ World Health Organization (WHO)

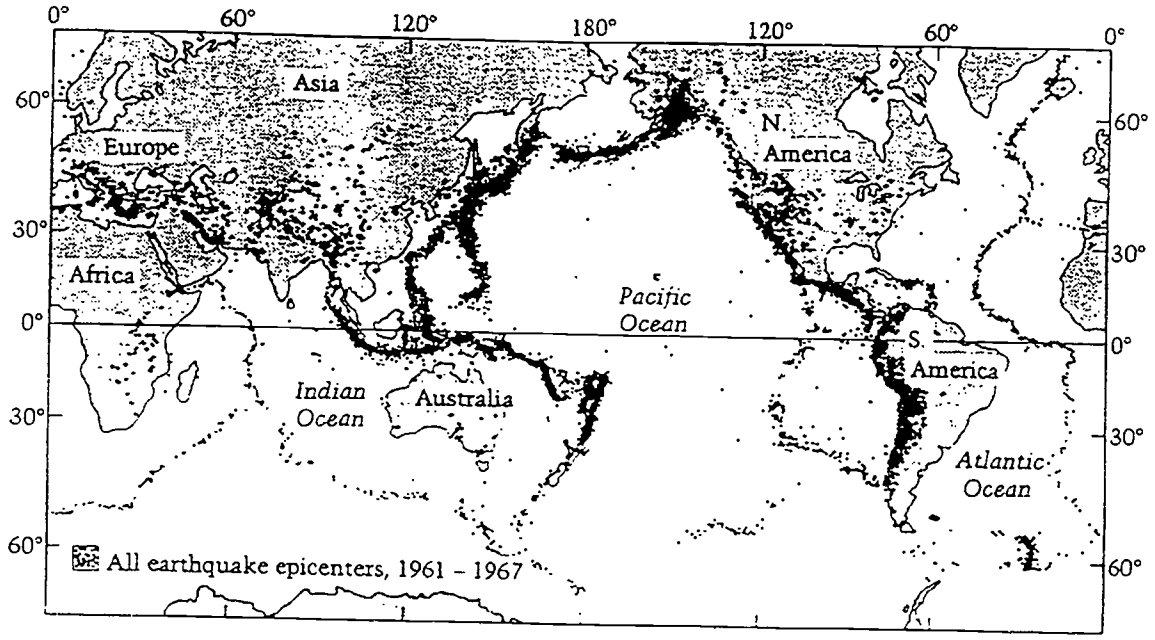
   Ⓒ North Atlantic Treaty Organization (NATO)

   Ⓓ British Commonwealth of Nations

K1000413

Grade Level:	12
Achievement Level:	Basic
Block Number:	Q3G4
Item Number:	3
Overall P-Value:	65
Avg Conditional P-Value:	62
Score Levels:	2

Questions 3–4 are based on the map below showing the distribution of earthquake epicenters around the world between 1961 and 1967.



4. What is responsible for the pattern of earthquake activity shown on the map?

Ⓐ Volcanic eruptions

Ⓑ The weight of ocean water pressing on the land

Ⓒ Hurricanes and cyclones

● The movement of tectonic plates

K100[illegible]91

Grade Level:	12
Achievement Level:	Basic
Block Number:	Q23G6
Item Number:	4
Overall P-Value:	75
Avg Conditional P-Value:	79
Score Levels:	2
Content Area:	1

Rationale Text:
Scoring Guide

Scoring Rationale: Student demonstrates ability to locate four features on a map showing Europe, Asia, Africa, and Australia.

4 - Complete. The response correctly labels all 4 features on the map.

3 - Essential. The response correctly labels 3 features on the map.

2 - Partial. The response correctly labels 1-2 feature on the map.

1 - Inappropriate. The response incorrectly labels all of the features on the map.

NB: If students write out answers instead of putting the numbers on the map, give credit for correct response.

Item Number: NONE Key: NONE Accession Number: KJ000479

Classification Codes:

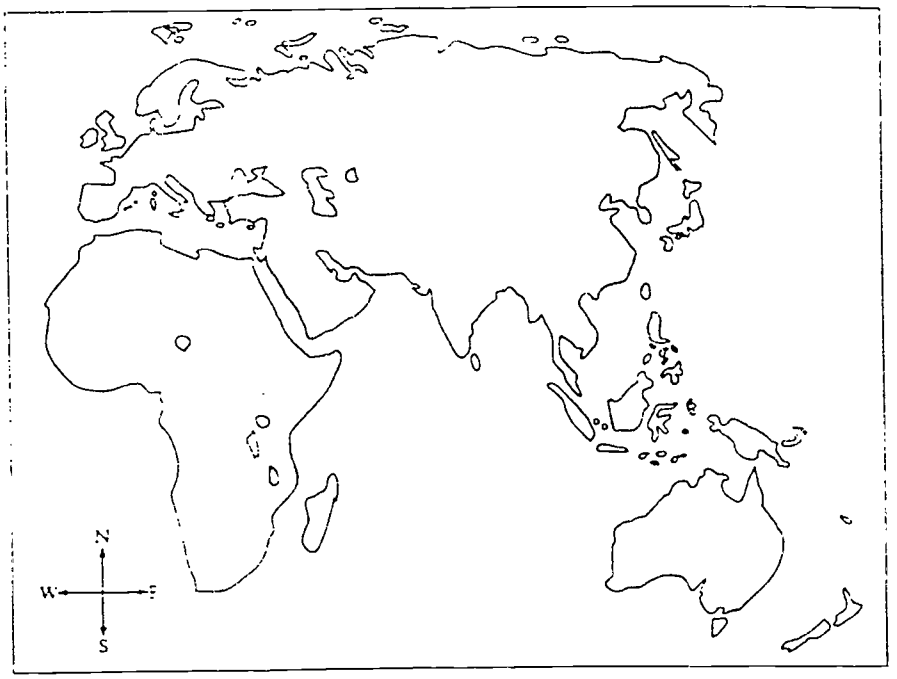
3	NA	NA	NA	NA	NA	TEXT

A dense canopy of interlaced broad-leafed trees blocks the sun and shades the forest floor all year in this sparsely populated area. Shallow-rooted trees that rise 150 feet or more have broad trunks with support roots above ground. Rain occurs every day and the forest floor is always damp and dark.

Rationale Text: NONE

12th-Grade Proficient—Students should have an extensive understanding and knowledge of the concepts and terminology of physical and human geography. They should be able to use geographic concepts to analyze spatial phenomena and to discuss economic, political, and social factors that define and interpret space. They should be able to do this through the interpretation of maps and other visual and technological tools, through the analysis of case studies, the utilization of data bases, and the selection of appropriate research materials. Students should be able to design their own maps based on descriptive data; describe the physical and cultural attributes of major world regions; relate the spatial distribution of population to economic and environmental factors; report both historical and contemporary events within a geographic framework using tools such as special purpose maps, and primary and secondary source materials.

For Question 5 write your answer on the map below.



5. Write the number of each of the following physical features in the correct location on the map on page 6

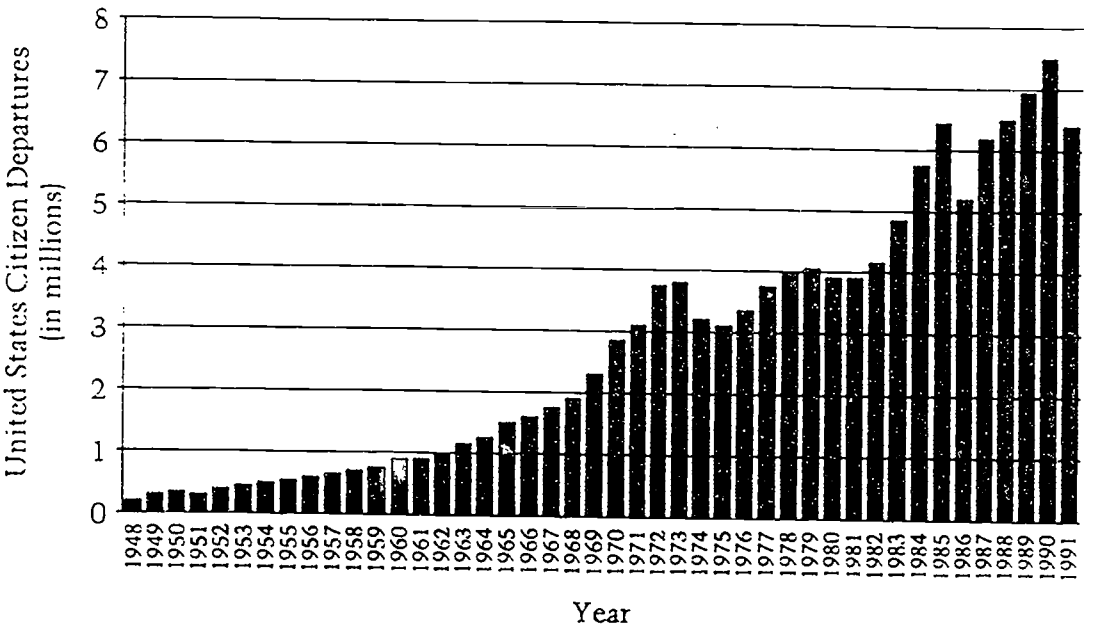
   1 Pyrenees Mountains
   2 The Japanese Archipelago
   3 Mediterranean Sea
   4 Persian Gulf

K1000486

Grade Level:	12
Achievement Level:	Proficient
Block Number:	Q3G4
Item Number:	5 ≥3
Overall P-Value:	23
Avg Conditional P-Value:	52
Score Levels:	4

Section 3

TRAVEL TO EUROPE, 1948-1991



9. What is the dominant trend shown in the graph?

Give two major reasons for the trend shown.

Grade Level:	12
Achievement Level:	Proficient
Block Number:	Q23G6
Item Number:	9 ≥3
Overall P-Value.	94
Avg Conditional P-Value:	55
Score Levels:	3
Content Area:	3

K1000520

Rationale Text:
Scoring Guide

Scoring Rationale: Student demonstrates ability to interpret what the dominant trend in a graph is and to give two reasons for the trend.

3 - **Complete.** The response identifies the major trend in the graph and provides two appropriate reasons to explain this.

2 - **Partial.** The response identifies the major trend in the graph but gives either no reason, or only one reason to explain this. Or, the response provides two appropriate reasons, but fails to identify the trend.

1 - **Inappropriate.** The response does not indicate the major trend shown in the graph, but may give one reason for the trend as suggested in the list given. It may just copy words from chart.

Credited Responses may include:

Trend in Graph

An increase in the number of United States citizens traveling to Europe between 1948 and 1991.

Appropriate Reasons

Improvements in travel	- air travel is more affordable
	- air travel is faster
	- air travel is safer
Increase in tourism	- more leisure time
	- more affluence
	- aging population
	- no major European way
	- cultural identity
	- more wanted to go
Internationalization of economy	- more business travel
	- more political travel

12th-Grade Advanced—Students should possess a comprehensive understanding of geographic knowledge and concepts; apply this knowledge to case studies; formulate hypotheses and test geographic models that demonstrate complex relationships between physical and human phenomena; apply a wide range of map skills; develop maps using fundamental cartographic principles including translating narratives about places and events into graphic representations, and use other visual and technological tools to perform locational analysis and interpret spatial relationships. Students should also be able to undertake sophisticated analysis from aerial photographs or satellite imagery and other visuals. Advanced students should be able to develop criteria assessing issues relating to human spatial organization and environmental stability and, through research skills and the application of critical thinking strategies, identify alternative solutions. They should be able to compile data bases from disparate pieces of information and from these data develop generalizations and speculations about outcomes when data change.

Mexico City is an example of runaway urban growth. Every day an estimated average of 1,700 people move there from villages in the countryside. In addition, more than 1,000 babies are born in the city daily. Some geographers think that as many as 50 million people will live there by the year 2000. Thousands of families survive on the equivalent of a few dollars a day, and most members of these families have no prospects for steady jobs or much improvement in the physical quality of their lives. However, regardless of the hardship and the poverty, people continue to pour into Mexico City. K1000401

4. Give two reasons why people continue to move to Mexico City despite the difficult living conditions. SE000757

Grade Level:	12
Achievement Level:	Advanced
Block Number	Q3G4
Item Number:	4 ≥3
Overall P-Value.	93
Avg Conditional P Value:	68
Score Levels:	3
Content Area:	3

Item Number: 4 Key: NONE Accession Number: SE000757

Classification Codes:
3 III C UNDER SA WRIT TEXT

Give two reasons why people continue to move to Mexico City despite the difficult living conditions.

Rationale Text:
Scoring Guide

Scoring Rationale: Student demonstrates ability to determine why people continue to move to Mexico City despite its difficult living conditions.

3 - Complete. The response gives two reasons why people continue to move to Mexico City. The idea must be conveyed that more opportunities not guarantees are available.

2 - Partial. The response gives one reason that explains why people continue to move to Mexico City.

1 - Inappropriate. The response does not give any reason explaining why people continue to move to Mexico City.

Credited Responses could include:

Reasons for Migrating to Mexico City

Opportunity to improve standard of living. Better housing, or education, or medical facilities, or employment opportunities

Difficulty of rural conditions. Escape poor housing, or medical facilities, or unemployment, or low wages, or poverty

Decline of traditional agriculture
Displacement of peasants
Relatives or friends in city - provide emotional and practical support.

Section 3

10. Fossil fuels such as oil and coal are formed from

Ⓐ geological processes that transform organic materials

Ⓑ the rapid decay of animal bones

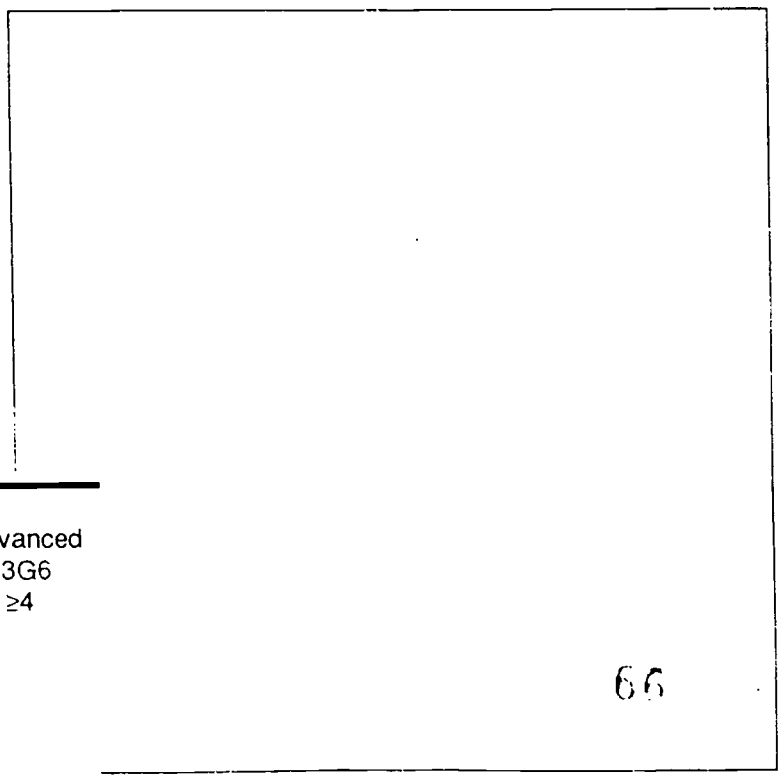
Ⓒ organic processes that lead to the fossilization of animal tissue

Ⓓ artificial processes used to treat and reuse garbage K10C0852

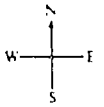
11. After we anchored our ships in the ocean and went ashore to explore, we marched west. The forest was so thick we could only travel three miles in the first two days. Then we came to the mountains and climbed to the top. A rushing river flowed west out of the mountains. We continued to march two miles west and came down out of the mountains. Two miles further we came to the coast. It was obvious that the area we were exploring was an isthmus.

In the box below, draw a map of the region described above. Be sure to include all of the geographical elements mentioned in the description. Include a scale to indicate distances.

K10C0849



Grade Level:	12
Achievement Level:	Advanced
Block Number:	Q23G6
Item Number:	11 ≥4
Overall P-Value:	86
Avg Conditional P-Value:	74
Score Levels:	4
Content Area:	1



Rationale Text:
Scoring Guide

Scoring Rationale: Student demonstrates an understanding of direction, isthmus, and simple map making.

4 - **Complete.** The response includes an accurate map in which at least 4 elements are correctly placed. The response must be an isthmus and have direction of travel and river correctly.

3 - **Essential.** The response includes a map in which 3 elements are correctly placed. The response may be a peninsula or an island.

2 - **Partial.** The response includes a map in which at least 2 elements are correctly placed.

1 - **Inappropriate.** The response does not include a map or the map shows none of the elements correctly.

NOTE: No answer that gets direction of map wrong can get more than a 2. Use of the scale is not necessary to get a 4.

Features on Map
East coast
Forest
Mountains
River flowing west
West coast

Figure 3

U.S. History Achievement Levels Descriptions and Exemplar Items

4th-Grade Basic

Fourth-grade students performing at the basic level should be able to identify and describe a few of the most familiar people, places, events, ideas, and documents in American history. They should be able to explain the reasons for celebrating most national holidays, have some familiarity with the geography of their own state and the United States, and be able to express in writing a few ideas about a familiar theme in American history.

14. Which area became part of the United States last?

● Hawaii

Ⓑ Texas

Ⓒ Oregon

Ⓓ Alaska

B000 143

Grade Level:	4
Achievement Level:	Basic
Block Number:	Q1H4
Item Number:	14
Overall P-Value:	55
Avg Conditional P-Value:	59
Score Levels:	2
Content Area:	4

Questions 11-12 refer to the statement below.

A house divided against itself cannot stand. I believe this government cannot endure permanently half slave and half free. I do not expect the Union to be dissolved—I do not expect the house to fall—but I do expect it will cease to be divided.

B0001108

12. What war broke out soon after the statement was made?

Ⓐ American Revolution

Ⓑ War of 1812

Ⓒ Mexican-American War

● Civil War

B0001110

Grade Level:	4
Achievement Level:	Basic
Block Number:	Q12H7
Item Number:	12
Overall P-Value:	54
Avg Conditional P-Value:	58
Score Levels:	2
Content Area:	1

4th-Grade Proficient

Fourth-grade students performing at the proficient level should be able to identify, describe and comment on the significance of many historical people, places, ideas, events, and documents. They should interpret information from a variety of sources, including texts, maps, pictures, and timelines. They should be able to construct a simple timeline from data. These students should recognize the role of invention and technological change in history. They should also recognize the ways in which geographic and environmental factors have influenced life and work.

7. Your teacher has asked you to teach your classmates about ONE of these famous people from American history:

George Washington
Sojourner Truth
Sitting Bull
Thomas Jefferson

Choose one of these famous people.

My famous person from American history is _____.

Write down three facts about the person you have chosen that would help you teach about that person.

Fact 1 _____

_____

Fact 2 _____

_____

Fact 3 _____

_____

BD001133

Grade Level:	4
Achievement Level:	Proficient
Block Number:	Q1H5
Item Number:	7 ≥3
Overall P-Value:	35
Avg Conditional P-Value:	65
Score Levels:	4
Content Area:	1

Q1H5_07

Item Number: 7 Accession Number: BO001133

Your teacher has asked you to teach you classmates about ONE of these famous piople from American history:

George Washington
Sojourner Truth
Sitting Bull
Thomas Jefferson

Choose one of these famous people.

My famous person from American history is _____.

Write down three facts about the person you have chosen that would help you teach about that person.

Fact 1 _____

_____

Fact 2 _____

_____

Fact 3 _____

_____

Rationale Text:

1 - **Inappropriate.** The response gives ones of the four names but is unable to provide any accurate facts about his or her identity, although some attempt is made to provide facts.

2 - **Partial.** The response gives one of the four names and provides one accurate fact about his/her identity.

3 - **Essential.** The response gives one of the four names and two accurate facts about his/her identity.

4 - **Complete.** The response gives one of the four names and three accurate facts abcut his/her identity.

NOTE: If no name is chosen, but accurate facts are given for one of the people listed in the item, the response may receive credit.

Also, facts derived from reading the question (for example, "my name is George") are unacceptable.

Credited Responses could include:

George Washington

Revolutionary War
president
had wooden teeth
lived in Mt. Vernon
survived Valley Forge winter
help his men to survive Valley Forge
crossed the Delaware
owned slaves
surveyor
Mt. Rushmore
on $1 bill
on quarter
helped to think of the plan that defeated the British at Yorktown
led Constitutional Convention
General in the French and Indian War (Sever-Years War)
fought in a war
General
do NOT accept the cherry tree story

Sojourner Truth

Famous African/American abolitionist
freed slave
spoke against slavery and for women's rights
helped to free slaves
speaker
Underground Railroad
gender and work rights
feminist
travelled through New England and Western States spreading her views

Sitting Bull

Native American
led his people in a war against the US gov't in the 1800's (1874) in order to prevent seizing of their lands & forced removal to reservations
Little Big Horn - Custer
Sioux/chief/Indian/medicine man
spent hears touring with a Wild West show
escaped with his people to Canada

Thomas Jefferson

writer of the Declaration of Independence
Secretary of State for Washington
president
sent Lewis and Clark on their expedition
had slaves
Ambassador to France
from Virginia
on the nickel
on the $2 bill
inventor - made machines (NOT light)

General Comments: NONE

8. What is the purpose of the Bill of Rights?

Ⓐ To say how much Americans should pay in taxes

● To protect freedoms like freedom of speech

Ⓒ To describe the jobs of the President and Congress

Ⓓ To make Washington, D.C., the capital of the United States

BOC3.131

Grade Level:	4
Achievement Level:	Proficient
Block Number:	Q1H5
Item Number:	8
Overall P-Value:	45
Avg Conditional P-Value:	65
Score Levels:	2
Content Area:	1

4. What is the main reason the Pilgrims and Puritans came to America?

● To practice their religion freely

Ⓑ To make more money and live a better life

Ⓒ To build a democratic government

Ⓓ To expand the lands controlled by the king of England

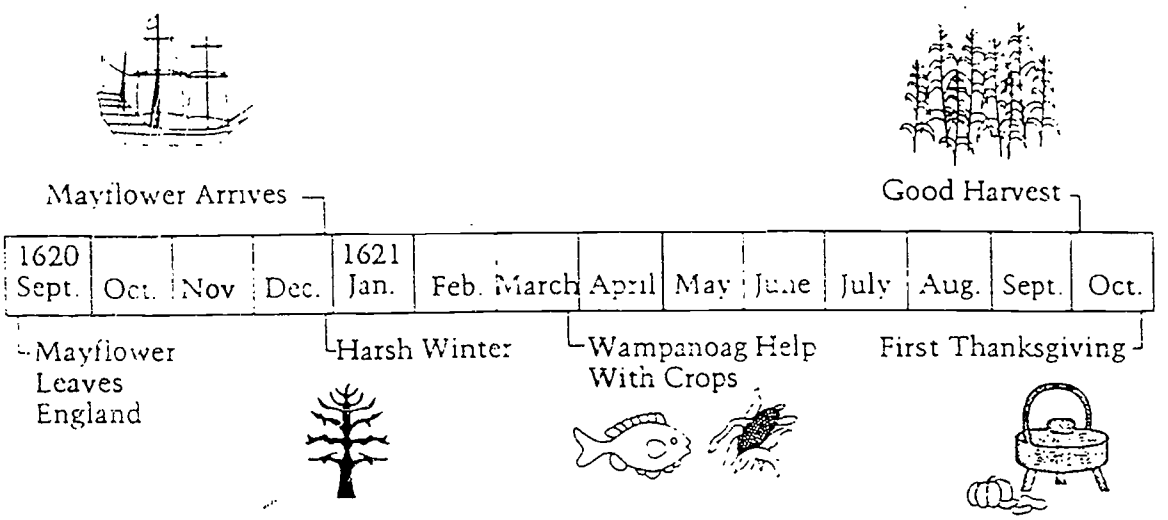
B0001005

Grade Level:	4
Achievement Level:	Proficient
Block Number:	Q12H7
Item Number:	4
Overall P-Value:	41
Avg Conditional P-Value:	80
Score Levels:	2
Content Area:	1

4th-Grade Advanced

Fourth-grade students performing at the advanced level should have a beginning understanding of the relationships between people, places, ideas, events and documents. They should know where to look for information, including reference books, maps, local museums, interviews with family and neighbors, and other sources. They should be able to use historical themes to organize and interpret historical topics, and to incorporate insights from beyond the classroom into their understanding of history. These students should understand and explain the role of invention and technological change in history. They should also understand and explain the ways in which geographic and environmental factors have influenced life and work.

Questions 3-4 refer to the time line below



3. What can you tell from the time line?

Ⓐ The Mayflower took more than one year to sail to Plymouth.

Ⓑ The Mayflower arrived in Plymouth in 1621.

● The first Thanksgiving was not celebrated in the same month as it is today.

Ⓓ The Pilgrims in Plymouth ate mostly fish.

BO000976

Grade Level:	4
Achievement Level:	Advanced
Block Number:	Q1H5
Item Number:	3
Overall P-Value:	32
Avg Conditional P-Value:	74
Score Levels:	2
Content Area:	1

2. Your teacher has asked you to teach your classmates about ONE of these famous places where an important event in American history happened:

the Alamo
Pearl Harbor
Gettysburg
Roanoke Island

My famous place in American history is _____

Write down three facts about the place that you have chosen that will help you teach your classmates about that place.

SC001135

Fact 1 _____

_____

Fact 2 _____

_____

Fact 3 _____

_____

Grade Level:	4
Achievement Level:	Advanced
Block Number:	Q12H7
Item Number:	2 ≥4
Overall P-Value	81
Avg Conditional P-Value:	62
Score Levels:	4
Content Area:	4

Item Number: 2 Accession Number: BO001136 Q12H7_C

Your teacher has asked you to teach your classmates about ONE of these famous places where an important event in American history happened:

the Alamo
Pearl Harbor
Gettysburg
Roanoke Island

My famous place in American history is _____

Write down three facts about the place that you have chosen that will help you teach your classmates about that place.

Fact 1 _____

_____

Fact 2 _____

_____

Fact 3 _____

_____

Rationale Text:
To determine whether students understand three important facts about one of the places listed.

1 - Appropriate. The answer gives no facts that are relevant to the particular place and that might help another person understand the place, but makes an attempt at giving facts.

2 - Partial. The answer gives one fact that is relevant to the particular place and that might help another person understand the place.

3 - Essential. The answer gives two facts that are relevant to the particular place and that would help another person understand the place.

4 - Complete. The answer gives three facts that are relevant to the particular place and that would help another person understand the place.

Q12H7_02

Credited responses could include :

Note - grade notations indicate information most likely to have been mastered in 4th or 8th grade)

the Alamo

(4th) - Spanish mission used as fort (in San Antonio, TX)
-Used as fort during battle fought by band of Texans demanding independence from Mexico.
-Texans defeated by Mexicans (March 1835)

(8th) - Texans fought (assisted by some Mexicans) under Colonel Travis
- Many Texans at Alamo were recent settlers from U.S.
- held out in fort, but defeated w/ barely survivors
- defeated by Mexicans led by General Santa Anna
- there was a war there
- *Davey Crockett, Jim Bowie died defending the Alamo, Sam Austin was not there*

Pearl Harbor (8th Grade)
- in Hawaii in the pacific
- bombed by the Japanese in 1941, (on December 7th)
- led to U.S. joining Allied forces in WWII - declared war against Japan
- most important navel base in Pacific
- *can accept many people died*

Gettysburg - (8th Grade)
- site of important battle in Civil War
- in Pennsylvania, *close to Washington DC*
- site where President Lincoln delivered important speech (in 1863) honoring soldiers who died in battle there
- President dedicated a part of the battlefield there to a cemetery for soldiers
- union forces defeated Confederate forces here, so badly that the battle was a turning point in the Civil War
- stopped Southern advance in Civil War

Roanoke Island (8th Grade)
- Off coast of present- day North Carolina
- Algonquin Indians lived there before white settlers arrived
- First English settlement there in 1585
- Soldiers led by Walter Raleigh est. a fort there in 1585
- soldiers there traded with Indians (1585), but returned home because of trouble with Indians and food shortages.
- second group of English settlers (led by John White) brought families, and received land, in order to make a permanent settlement in 1587
- this second group of settlers disappeared

General Comments: NONE

8th-Grade Basic

Eighth-grade students performing at the basic level should be able to identify and place in context a range of historical people, places, events, ideas, and documents. They should be able to distinguish between primary and secondary sources. They should have a beginning understanding of the diversity of the American people and the ways in which people from a wide variety of national and cultural heritages have become part of a single nation. Eighth-grade students at the basic level should also have a beginning understanding of the fundamental political ideas and institutions of American life and their historical origins. They should be able to explain the significance of some major historical events.

Questions 4-6 refer to the passage below.

We hold these truths to be self-evident: That all men are created equal, that they are endowed by their Creator with certain unalienable rights, that among these are life, liberty, and the pursuit of happiness. That, to secure these rights, governments are instituted among men, deriving their just powers from the consent of the governed, that, whenever any form of government becomes destructive of these ends, it is the right of the people to alter or to abolish it, and to institute a new government.

—1776

BOOC234

5. The primary author of the document was

Ⓐ George Washington

Ⓑ John Marshall

Ⓒ Robert E. Lee

● Thomas Jefferson

BOOC1234

Grade Level:	8
Achievement Level:	Basic
Block Number:	Q2H4
Item Number:	5
Overall P-Value:	71
Avg Conditional P-Value:	79
Score Levels:	2
Content Area:	1

Questions 11-12 refer to the statement below.

A house divided against itself cannot stand. I believe this government cannot endure permanently half slave and half free I do not expect the Union to be dissolved—I do not expect the house to fall—but I do expect it will cease to be divided.

BO001108

11. The statement was made by

Ⓐ George Washington

Ⓑ Thomas Jefferson

● Abraham Lincoln

Ⓓ Theodore Roosevelt

BOC011C?

Grade Level:	8
Achievement Level:	Basic
Block Number:	Q12H7
Item Number:	11
Overall P-Value:	55
Avg Conditional P-Value.	70
Score Levels:	2
Content Area:	1

1. During the 1500's and 1600's, what was the major cause of death among Indians of the Americas?

   Ⓐ Warfare among tribes

   Ⓑ Warfare between Native Americans and Europeans

   ● Infections and diseases brought by Europeans

   Ⓓ Changing climatic conditions

Grade Level:	8
Achievement Level:	Basic
Block Number:	Q23H9
Item Number:	1
Overall P-Value:	69
Avg Conditional P-Value:	75
Score Levels:	2
Content Area:	2

8th-Grade Proficient

Eighth-grade students performing at the proficient level should be able to explain the significance of people, places, events, ideas, and documents, and to recognize the connection between people and events within historical contexts. They should understand and be able to explain the opportunities, perspectives and challenges associated with a diverse cultural population. They should incorporate geographic, technological, and other considerations in their understanding of events and should have knowledge of significant political ideas and institutions. They should be able to communicate ideas about historical themes while citing evidence from primary and secondary sources to support their conclusions.

10. What was the main reason that many leaders in Great Britain leaned toward supporting the Confederacy in the Civil War?

Ⓐ Plantation owners in Britain held slaves

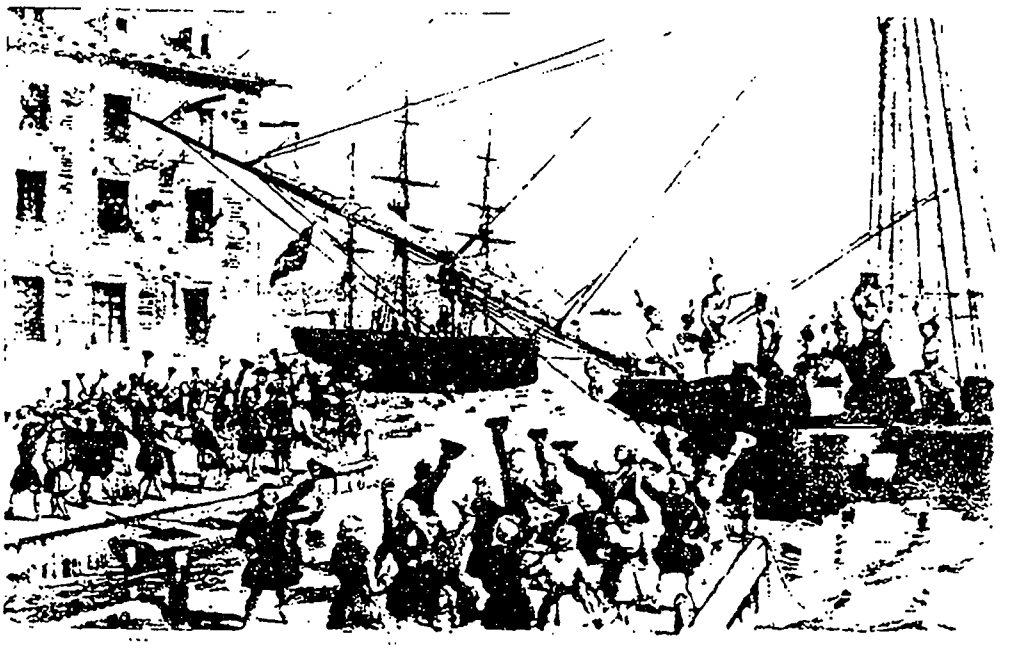
Ⓑ Most British immigrants to the United States lived in the South

● Britain relied on Southern cotton for its factories.

Ⓒ British politicians wanted to make the United States weaker so that they could conquer it.

B0C01123

Grade Level:	8
Achievement Level:	Proficient
Block Number:	Q23H9
Item Number:	10
Overall P-Value:	41
Avg Conditional P-Value:	63
Score Levels:	2
Content Area:	4



The Destruction of Tea at Boston Harbor N. Currier 1846. Museum of the City of New York. The Harry T Peters Collection

3. Identify the event that is portrayed in the picture above.

Why is the event important in United States history? B0001902

Grade Level:	8
Achievement Level:	Proficient
Block Number:	Q2H4
Item Number:	3 ≥3
Overall P-Value:	97
Avg Conditional P-Value:	55
Score Levels:	3
Content Area:	1

Item Number: 3 Accession Number BO001902
Q2H4_03

- insert picture of Boston Tea Party -

Identify the event that is portrayed in the picture above.

_____

_____

_____

Why is the event important in United States history?

_____

_____

_____

Rationale Text:
Students can identify the Boston Tea party from an engraving of the event and explain something about the event's significance.

1 = **Inappropriate.** The response does not correctly identify the event, or offer any information relevant to the event.

2 = **Partial.** The response correctly identifies the event (as the Boston Tea Party and does NOT just repeat the title of the picture) or explains what the colonists are doing without addressing its historical significance. OR, it describes the historical significance but does not explicitly identify the event as the Boston Tea Party.

3 = **Appropriate.** The response correctly identifies the event as the Boston Tea Party. It also explains that the event is important because it showed colonial determination to combat undesirable British policies, and to seize control of the colonial economy and of colonial politics.

General Comments: NONE

3's	2's	1's
Identifies AND Explains	Identifies OR Explains	Neither identifies nor explains
• ID: Boston Tea Party engraving description		repeats title of
• Explanation step on road to revolution political activity result of frustration regional importance - hotbed of resistance		

8th-Grade Advanced

Eighth-grade students performing at the advanced level should recognize significant themes and movements in history and begin to understand particular events in light of these themes and movements. They should have an awareness of continuity and change over time and be able to draw relevant analogies between past events and present-day situations. They should be able to frame questions about historical topics and use multiple sources to develop historical generalizations and interpretations. They should be able to explain the importance of historical themes, including some awareness of their political, social, and economic dimensions.

6. What was one consequence of Nat Turner s rebellion?

Ⓐ Large numbers of slaves fled to the North.

Ⓑ Slave revolts broke out throughout the South.

Ⓒ Conditions for slaves on many southern plantations improved.

● Southern states passed laws designed to tightly control slaves

BO061543

Grade Level:	8
Achievement Level:	Advanced
Block Number:	Q23H9
Item Number:	6
Overall P-Value:	26
Avg Conditional P-Value	62
Score Levels:	2
Content Area:	1

14. What goal was most important in shaping United States foreign policy between 1945 and 1990 ?

● Preventing the spread of communism to new areas and weakening it where it already existed

Ⓑ Encouraging trained scientists and other skilled workers who lived in foreign countries to immigrate to the United States

Ⓒ Strengthening the United States industrial and agricultural sectors to help them compete against the British and the French

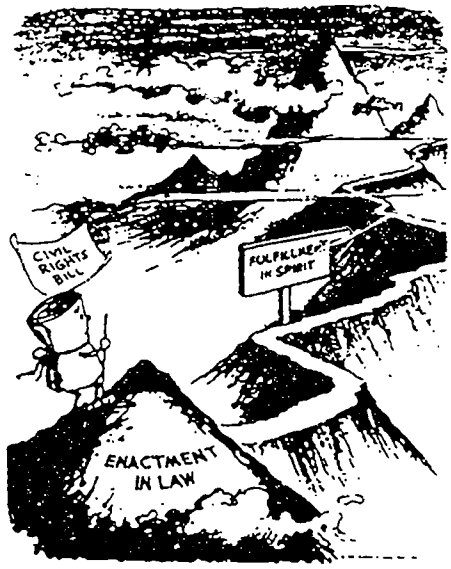
Ⓓ Providing foreign aid to all poor countries to help them develop economically and technologically

BO001490

Grade Level:	8
Achievement Level:	Advanced
Block Number:	Q23H9
Item Number:	14
Overall P-Value:	26
Avg Conditional P-Value:	85
Score Levels:	2
Content Area:	4

Questions 16-17 refer to the cartoon below.

BC001462



There's a long, long trail a-winding

Bruce Shanks in the Buffalo News.

17. What is the main message of this cartoon?

BO001463

Grade Level:	8
Achievement Level:	Advanced
Block Number:	Q23H9
Item Number:	17 ≥2
Overall P-Value:	11
Avg Conditional P-Value:	69
Score Levels:	3
Content Area:	1

Item Number: 17 Accession Number: B0001464 Q23119

What is the main message of this cartoon?

insert 5 lines

Rationale Text:

1 - **Inappropriate.** The response does not correctly identify the message of the cartoon. It does not indicate that the cartoon refers to problems in the civil rights movement in the United States.

2 - **Partial.** The response understands that the cartoon is about problems associated with the civil rights bill but is not able to explain the cartoon in a specific and precise fashion.

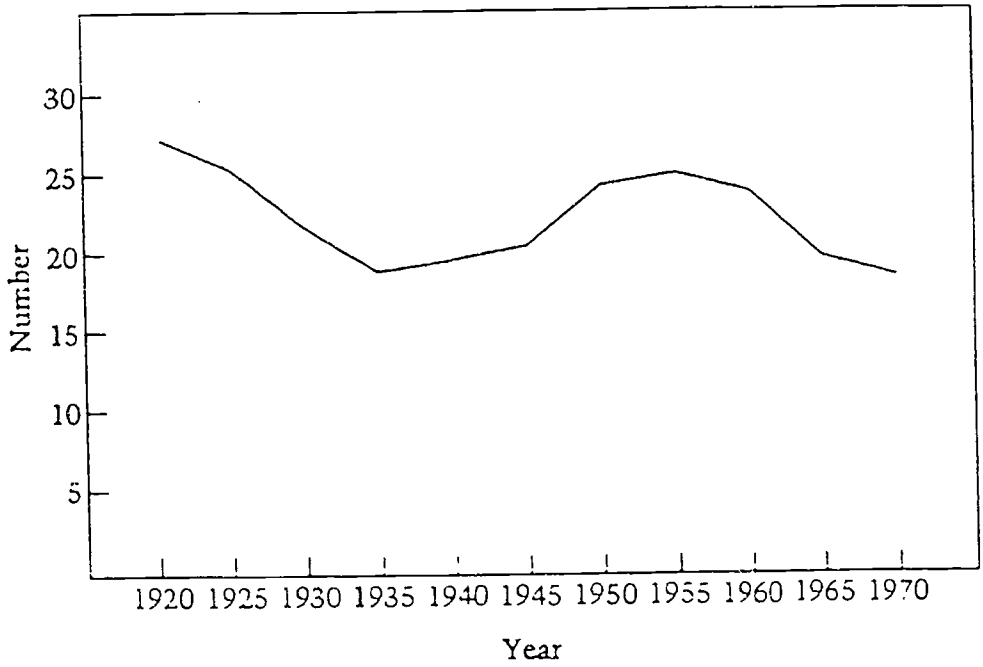
3 **Appropriate.** The response correctly identifies the main message of the cartoon as being that passing civil rights laws does not mean that civil rights issues are resolved, and that social, economic, and political (and not just legal) changes were also necessary.

General Comments: NONE

12th-Grade Basic

Twelfth-grade students performing at the basic level should be able to identify the significance of many people, places, events, dates, ideas, and documents in U.S. history. They should also recognize the importance of unity and diversity in the social and cultural history of the United States, and an awareness of America's changing relationships with the rest of the world. They should have a sense of continuity and change in history and be able to relate relevant experience from the past to their understanding of contemporary issues. They should recognize that history is subject to interpretation and should understand the role of evidence in making an historical argument.

LIVE BIRTHS PER 1,000 RESIDENTS IN THE UNITED STATES



12. The graph above supports which statement about the birthrate in the United States?

Ⓐ It declined steadily from 1920 to 1950.

Ⓑ It increased rapidly during the Great Depression.

● It increased after the Second World War.

Ⓓ It tended to increase after each war.

BO001573

Grade Level:	12
Achievement Level:	Basic
Block Number:	Q3H5
Item Number:	12
Overall P-Value:	49
Avg Conditional P-Value:	60
Score Levels:	2
Content Area:	2

7. The Monroe Doctrine was intended to

Ⓐ promote United States trade with China

Ⓑ help keep the peace in Europe

● discourage European involvement in the Americas

Ⓓ protect United States business in Japan and Korea

B0001452

Grade Level:	12
Achievement Level:	Basic
Block Number:	Q23H9
Item Number:	7
Overall P-Value:	41
Avg Conditional P-Value:	50
Score Levels:	2
Content Area:	4

12th-Grade Proficient

Twelfth-grade students performing at the proficient level should understand particular people, places, events, ideas, and documents in historical context, with some awareness of the political, economic, geographic, social, religious, technological, and ideological factors that shape historical settings. They should be able to communicate reasoned interpretations of past events, using historical evidence effectively to support their positions. Their written arguments should reflect some in-depth grasp of issues and refer to both primary and secondary sources.

"Our reconstruction measures were radically defective because they failed to give the ex-slaves any land."

—Frederick Douglass

4. Describe briefly the way in which Douglass's statement helps explain the rise of sharecropping in the South after the Civil War. In your answer, be sure to define the term sharecropping. B00C173C

Grade Level:	12
Achievement Level:	Proficient
Block Number:	Q3H5
Item Number:	4 ≥3
Overall P-Value:	19
Avg Conditional P-Value:	57
Score Levels:	4
Content Area:	1

Item Number: 4 Accession Number: BO001730 Q3H5_04

"Our reconstruction measures were radically defective because they failed to give the ex-slaves any land."

-- Frederick Douglass

Describe briefly the way in which Douglass's statement helps explain the rise of sharecropping in the South after the Civil War. In your answer, be sure to define the term sharecropping.

Rationale Text:
To determine whether students understand what sharecropping was and can link it to reconstruction measures.

1 - **Inappropriate.** The response does not understand sharcropping or Reconstruction.

2 - **Partial.** The response is able to define sharecropping or is able to show an understanding of what sharecropping is, but does not explain its use or its rise.

3 - **Essential.** The response accurately defines sharecropping. The response links the use of sharecropping to general economic conditions faced by ex-slaves (causal), but does not link the use of sharecropping to the failures of reconstruction policy in general or to the failure of reconstruction to bring about land reform. For example, "slaves had it rough/poor."

4 - **Complete.** The response accurately defines sharecropping. The response explains the link between the rise of sharecropping and specific reconstruction policies or explains the specific ways in which the failure to give ex-slaves land created a cycle (of poverty) that led to sharecropping. For example: "no mobility or stuck and couldn't get North."

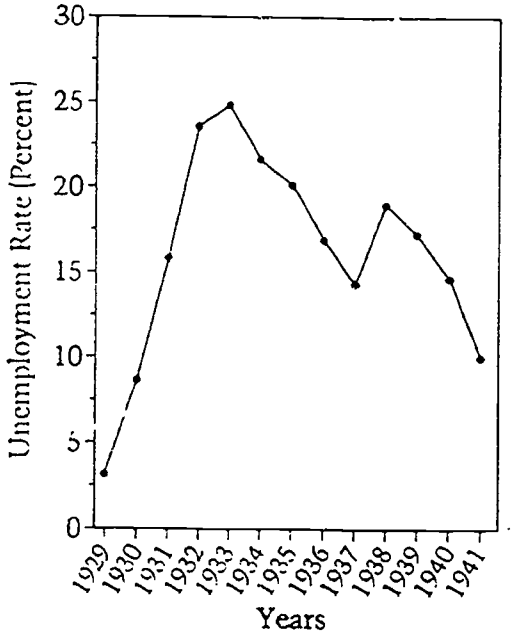
Definition of Sharecropping (for a '2', can be implicit)
- tax in kind
- storeowner "loans"
- racism
- lack of ownership
- description of how sharecropping came to be according to Douglass

General Comments: NONE

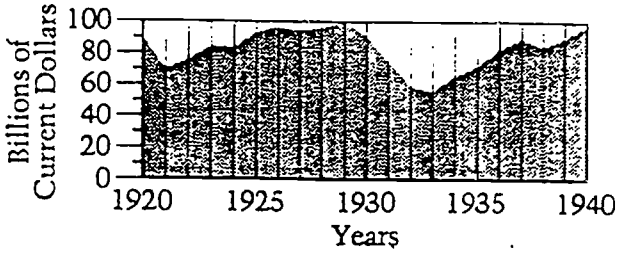
Document E – *Charts and graphs showing economic information about the Great Depression*

BO001855

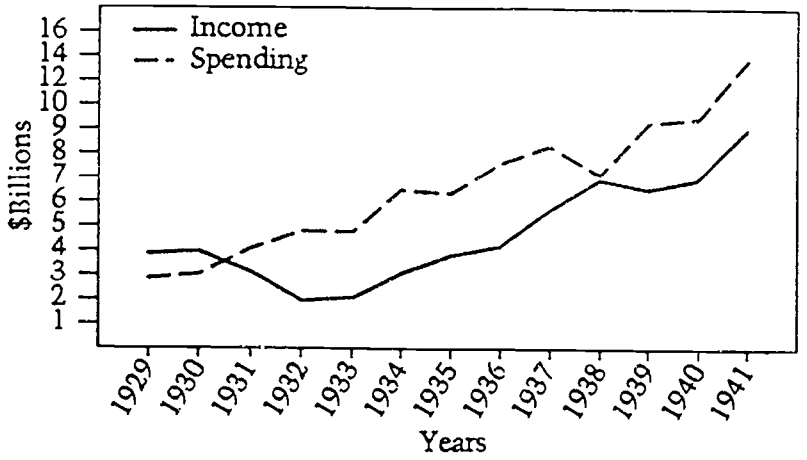
UNEMPLOYMENT RATES, 1929–1941



GROSS NATIONAL PRODUCT, 1920–1940



THE FEDERAL BUDGET, 1929–1941



6. Based on the graphs and your knowledge of history, you can conclude that the first successful steps toward partial recovery from the Great Depression occurred

Ⓐ at the end of the First World War

Ⓑ with the passage of the Hawley-Smoot Tariff

Ⓒ with United States entry into the Second World War

● with the election of Franklin Delano Roosevelt as President

BO001856

Grade Level:	12
Achievement Level:	Proficient
Block Number:	Q3H11
Item Number:	6
Overall P-Value:	41
Avg Conditional P-Value:	64
Score Levels:	2
Content Area:	3

12th-Grade Advanced

Twelfth-grade students achieving at the advanced level should demonstrate a comprehensive understanding of events and sources of U.S. history. Recognizing that history is subject to interpretation, they should be able to evaluate historical claims critically in light of the evidence. They should understand that important issues and themes have been addressed differently at different times and that America's political, social, and cultural traditions have changed over time. They should be able to write well-reasoned arguments on complex historical topics and draw upon a wide range of sources to inform their conclusions.

17. What is the main message of this cartoon? BOOC1464

Grade Level	12
Achievement Level	Advanced
Block Number	Q23H9
Item Number	17 ≥3
Overall P-Value	87
Avg Conditional P Value	75
Score Levels	3
Content Area	1

Q23119

Item Number: 17 Accession Number: BO001464

What is the main message of this cartoon?

insert 5 lines

Rationale Text:

1 - **Inappropriate.** The response does not correctly identify the message of the cartoon. It does not indicate that the cartoon refers to problems in the civil rights movement in the United States.

2 - **Partial.** The response understands that the cartoon is about problems associated with the civil rights bill but is not able to explain the cartoon in a specific and precise fashion.

3 **Appropriate.** The response correctly identifies the main message of the cartoon as being that passing civil rights laws does not mean that civil rights issues are resolved, and that social, economic, and political (and not just legal) changes were also necessary.

General Comments: NONE

Document F – *Two people's memories of the early years of the Depression*

Oh, I remember having to move out of our house. My father had brought in a team of horses and wagon. We had always lived in that house, and we couldn't understand why we were moving out. When we got the other house, it was a worse house, a poor house. That must have been around 1934. I was about six years old.

— Cesar Chavez

The oil boom come [to Oklahoma City] in '29. People come from every direction in there. A coupla years later, they was livin' in everything from pup tents, houses built out of cardboard boxes and old pieces of metal that they'd pick up—anything that they could find to put somethin' together to put a wall around 'em to protect 'em from the public.

— Mary Owsley

9. What effect did the Depression have on families? Use both the passages and your knowledge of history in your answer.

BO001859

Grade Level	12
Achievement Level	Advanced
Block Number	Q3H11
Item Number	9 [illegible]3
Overall P-Value	96
Avg Conditional P-Value	76
Score Levels	3
Content Area	3

Item Number: 9 Accession Number: BO001859 Q3H11_09

Document F.- *Two people's memories of the early years of the Depression.*

Oh, I remember having to move out of our house. My father had brought in a team of horses and wagon. We had always lived in that hours, and we couldn't understand why we were moving out. When we got the other house, it was a worse house, a poor house. That must have been around 1934. I was about six years old.

- Cesar Chaves

The oil boom cam (to Oklahoma City) in '29. People come from every direction in there. A couple years later, they was livin' in everything from pup tents, houses built out of cardboard boxes and old pieces of metal they'd pick up -- anything that they could find to put somethin' together to put a wall around 'em to protect 'em from the public.

- Mary Owsley

What effect did the Depression have on families? Use both the passages and your knowledge of history in your answer.

Rationale Text:

1. - **Inappropriate.** The response does not indicate that the Depression led many families to become homeless or live in bad conditions, etc. OR, the responses are significantly exaggerated or inaccurate.

2. - **Partial.** The response indicates that the Depression is associated with homelessness and led people to live in bad conditions. It does not use evidence in the passage or other historical data for support. OR, the response uses passages correctly but very leterally fails to generalize from them.

3. - **Appropriate.** The response indicates that the depressions dislocated and made many families homeless. The answer may also indicate that the Depression put huge tensions on many families. The answer refers to the passage or historical evidence in the response. The response shows a sense of the specific changes families were forced to undergo, and not just of bad conditions.

General Comments: NONE

Figure 4a
Understanding of Student Performance: Geography
5 = Very Well Formed; 1 = Not Well Formed

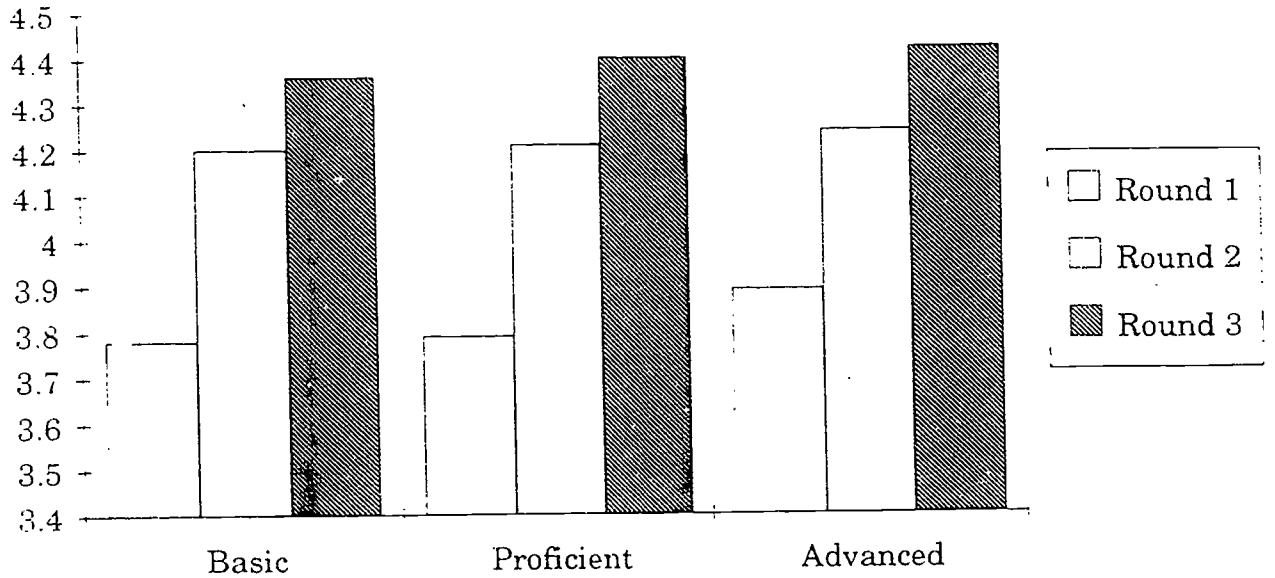


Figure 4b
Understanding of Student Performance: History
5 = Very Well Formed; 1 = Not Well Formed

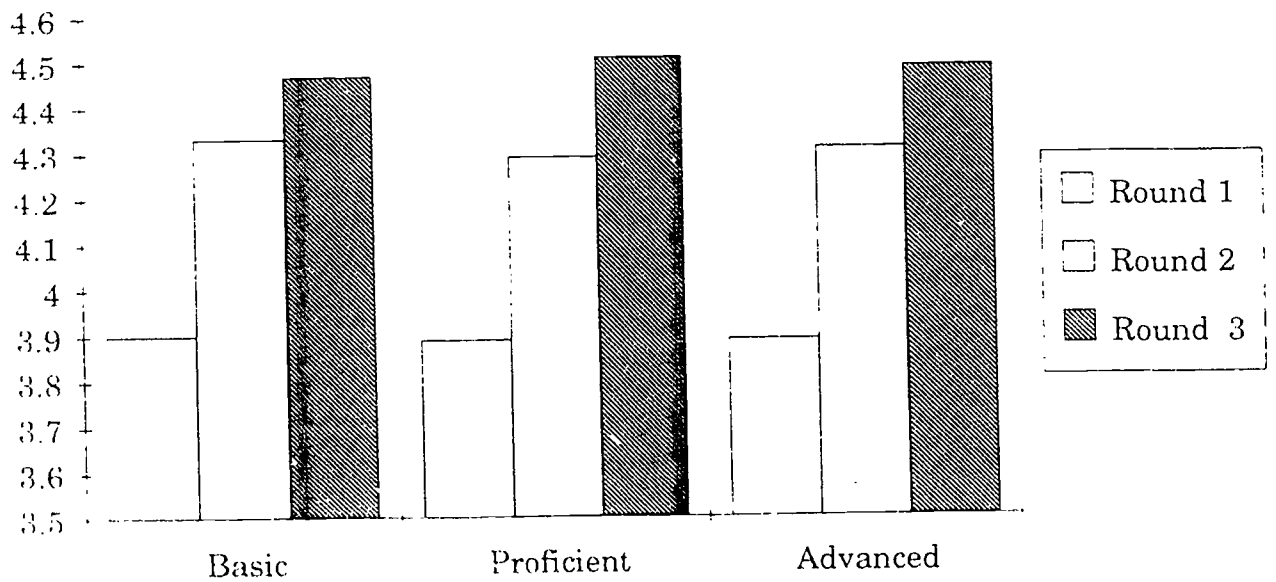


Figure 5a
Conceptualization of Student Performance: Geography
5 = Very Well Formed; 1 = Not Well Formed

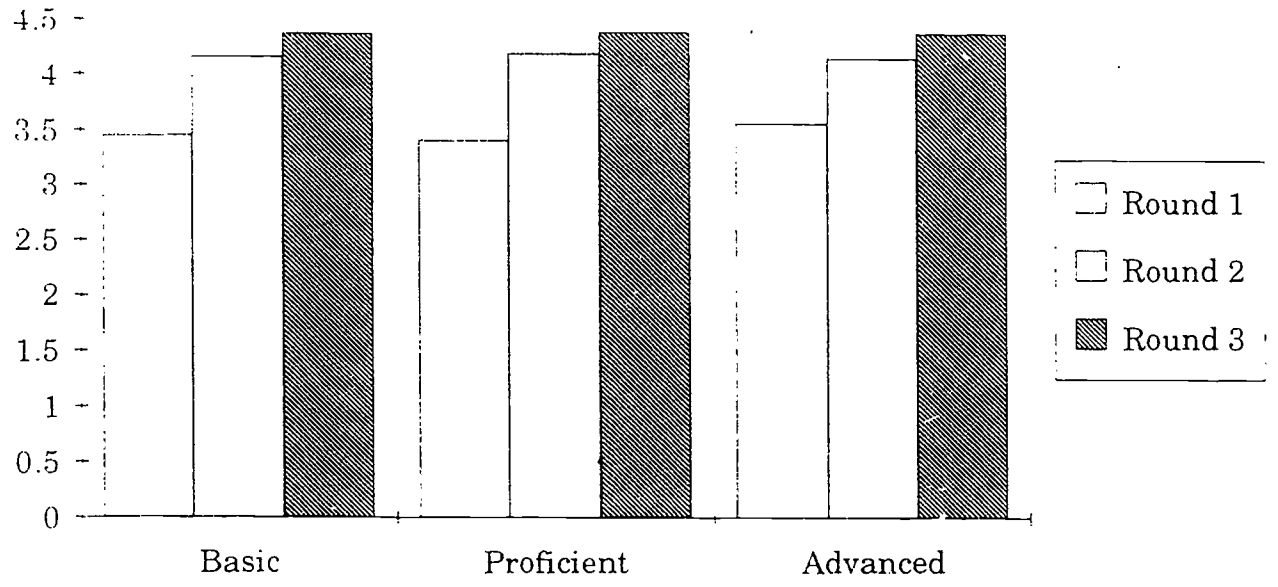


Figure 5b
Conceptualization of Student Performance: History
5 = Very Well Formed; 1 = Not Well Formed

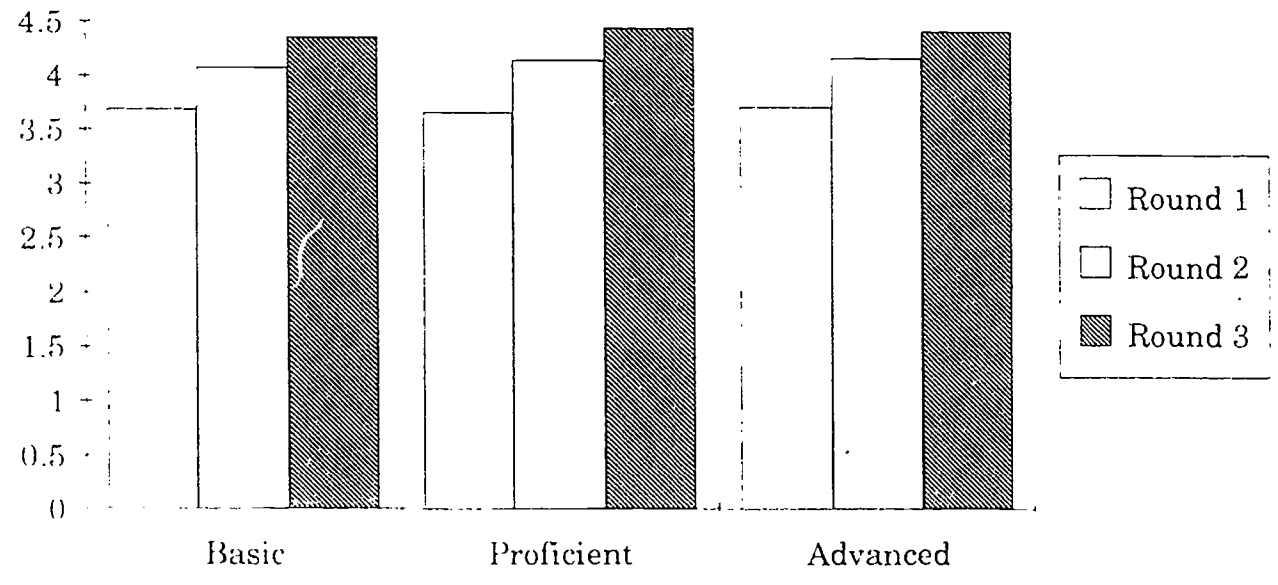


Figure 6a
Geography: Clarity of Rating Methods, by Rounds

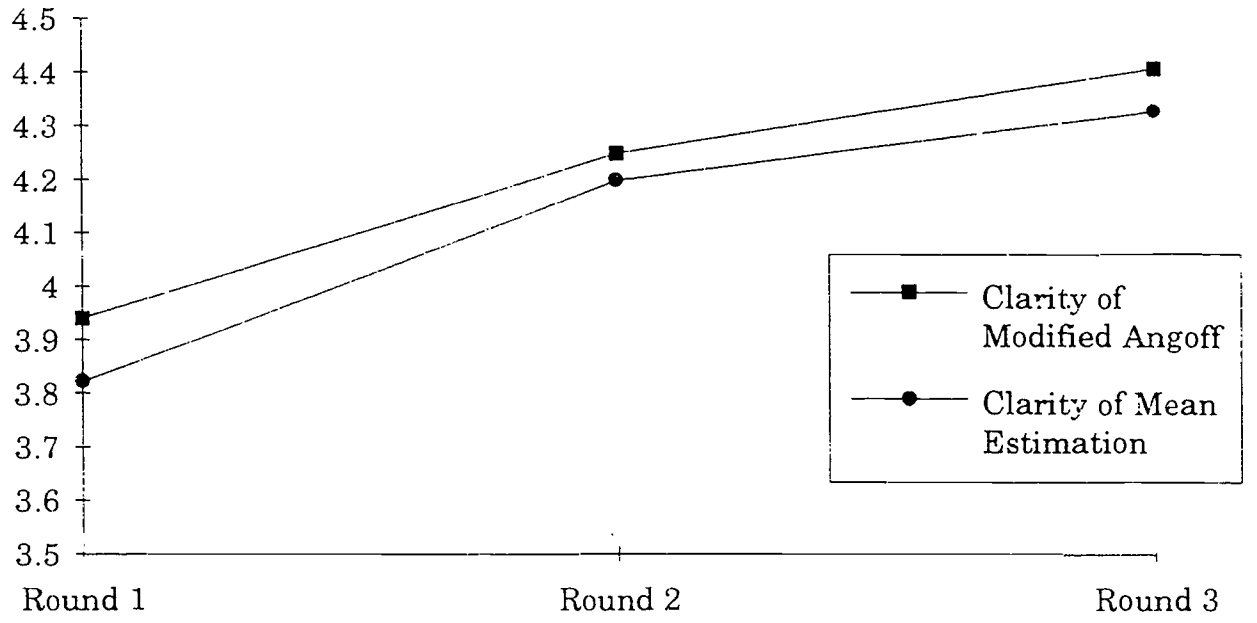


Figure 6b
History: Clarity of Rating Methods, by Rounds

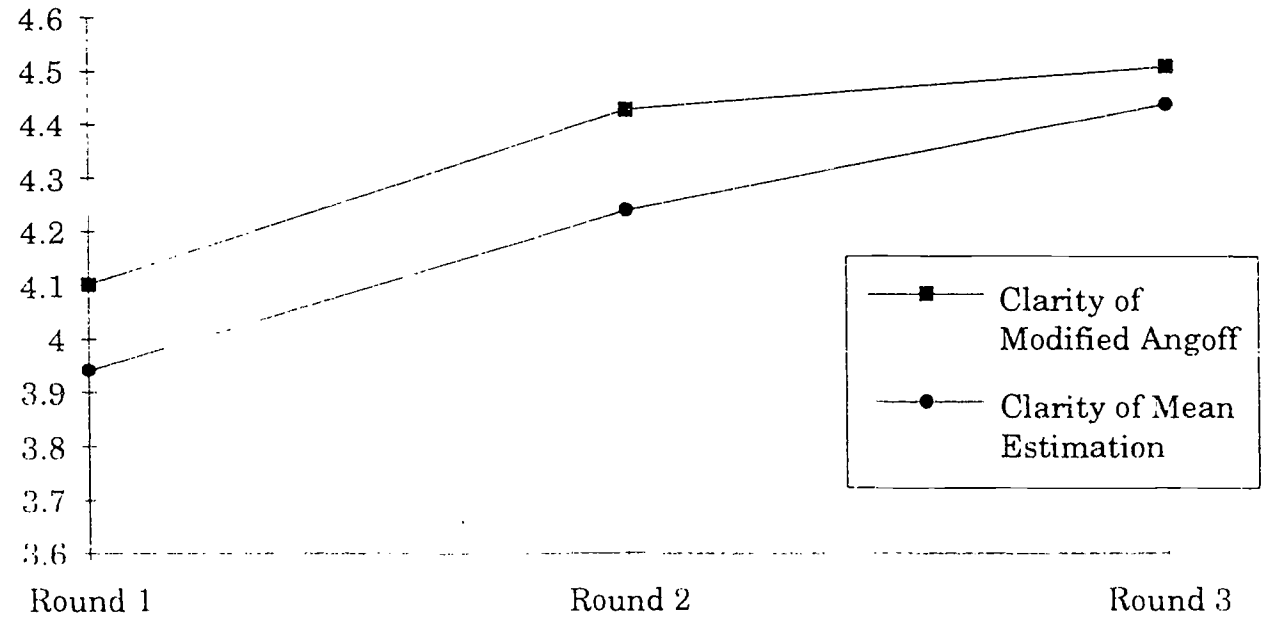


Figure 7a
Geography: Ease of Applying Rating Methods, by Rounds

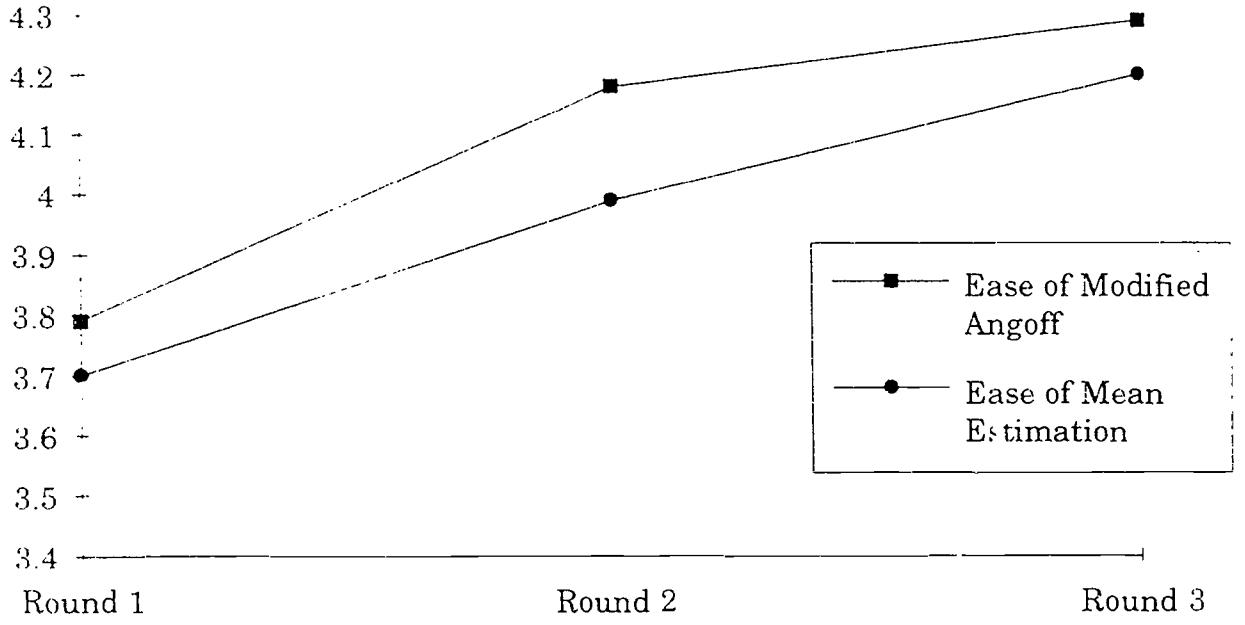


Figure 7b
History: Ease of Applying Rating Methods, by Rounds