American Students' Perspectives on Alternative Assessment: Do They Know It's Different?

CSE Technical Report 439

Joan L. Herman, Davina C. D. Klein, and Sara T. Wakai
CRESST/University of California, Los Angeles

July 1997

National Center for Research on Evaluation,
Standards, and Student Testing (CRESST)
Center for the Study of Evaluation (CSE)
Graduate School of Education & Information Studies
University of California, Los Angeles
Los Angeles, CA 90024-6511
(310) 206-1532

The work reported herein was supported under the Educational Research and Development Center Program cooperative agreement R117G10027 and CFDA Catalog number 84.117G as administered by the Office of Educational Research and Improvement, U.S. Department of Education.

AMERICAN STUDENTS' PERSPECTIVES ON ALTERNATIVE ASSESSMENT: DO THEY KNOW IT'S DIFFERENT?[1]

Joan L. Herman, Davina C. D. Klein, and Sara T. Wakai

CRESST/University of California, Los Angeles

Abstract

This study used the 1993 California Learning Assessment System (CLAS) Middle Grades Mathematics Performance Assessment as a platform to examine alternative assessment in actual practice in the U.S. Reported here is information gathered using the CLAS regarding student attitudes and approaches toward this new type of assessment. At issue is whether students find alternative assessments to be more motivating and interesting than traditional types of tests, and whether they appreciate the difference between traditional and alternative tasks. Data were collected in 13 schools across the state of California, involving more than 800 students. Instrumentation used in data collection included student surveys and in-depth student retrospective interviews. Findings suggest that students do indeed understand the differences in approaches necessitated by novel open-ended tasks versus more familiar multiple-choice tasks. In addition, student attitudes toward these two types of tasks are discussed in detail.

Educational policy makers at the national, state, and local levels in the United States continue to act on their beliefs in the power of educational assessment to improve schools. Through new mandated assessments, policy makers believe they can communicate standards; motivate and monitor progress toward attainment of those standards; provide useful feedback to all in the school community; and hold schools, and the teachers and students within them, accountable for improved performance. Their beliefs are bolstered by research showing that traditional testing has encouraged teachers and students to focus on what is tested (Herman & Golan, 1991; Madaus, 1991; Shepard, 1991). Unfortunately, due to the test content on traditional, standardized tests, this

[1] The authors wish to thank Steve D'Amico for his assistance, and all the teachers and students for their participation in this study.

teaching-to-the-test has resulted in a distortion of the curriculum for many students, narrowing it to basic, low-level skills (Herman & Dorr-Bremme, 1983; Herman & Golan, 1991; Kellaghan & Madaus, 1991; Shepard, 1991; Smith & Rottenberg, 1991). The result: Teachers, administrators, and policy makers across the United States are seeking new kinds of assessments whose content will reflect rigorous standards for student accomplishment. These new assessments will encourage schools to teach and students to learn the complex knowledge and problem-solving skills needed for future success.

Unlike traditional tests, new alternative assessments encourage students to think critically and draw their own conclusions to complex problems. Rather than asking students to select answers to short, discrete questions—often devoid of real-world context or application—these new assessments invite students to create extended responses, using multiple modes of representation. New assessments minimize the importance of rigid time constraints; they also encourage students to use tools (such as calculators) to help them in solving the novel problems on the assessments. Students' responses to real-life, "authentic" problems are scored by educators exercising judgment, not by machines reading "bubbles"; students' thinking processes, as well as their products, are often taken into consideration in the scoring rubrics.

For large-scale assessment purposes in the United States, as well as in international assessments, alternative assessment typically means assessment tasks that each ask students to create a response over a time span ranging from 20 minutes to a few classroom periods. For example, tasks might include having students design a bookcase within given function, space, and cost constraints and then asking them to explain how their design meets the given parameters (California State Department of Education, 1992); asking students to study the motion of maple seeds, to design experiments to explain their spinning flight patterns, and to interpret results in terms of scientific concepts such as laws of motion, aerodynamics, and air resistance (Lomask, Baron, Greig, & Harrison, 1992); prompting students to write a letter in French comparing the benefits of living in the country and living in a big city (Carroll, 1975); asking students to read historical documents and then use the perspectives in these documents with their prior knowledge to explain a major historical issue to a peer (Baker, Freeman, & Clayton, 1991); and asking students to make a presentation on their proposal for disposing of nuclear waste, based on a semester of work and their knowledge of science, and taking into consideration social, political, and environmental issues (Herman, Osmundson, & Pascal, 1996).

Claims for these kinds of assessments are frequent in the literature and on national conference agendas in the U.S. One claim is that students will find these assessments more meaningful and more motivating than traditional tests. These claims stem from the realistic and complex nature of the problems: *All* students should be encouraged by these types of tasks to show what they know and can do, rather than just those students who are motivated by the external rewards afforded them in high standardized test scores. Another claim is that these assessments truly stimulate students to engage in complex thinking and thus reflect higher standards of excellence than old-style standardized tests. Their ability to target higher level thinking and problem-solving skills makes these assessments suitable targets for instruction.

While the rhetoric is abundant, due to the relative newness of alternative assessments, evidence substantiating the above claims is just beginning to be accumulated. The study reported here is a portion of a larger investigation in which we used the 1993 California Learning Assessment System (CLAS) Middle Grades Mathematics Performance Assessment as a platform to examine alternative assessment in actual practice (Herman, Klein, Heath, & Wakai, 1995). In our study, we gathered information about student attitudes toward this new type of assessment. Do students really find alternative assessments more interesting and more motivating than traditional types of assessments? Do they understand and appreciate the difference between traditional and alternative tasks, both in terms of what is required of them as students and in terms of what information their teachers will garner from their responses to these assessments? Because the literature has shown subgroup differences in students' performance and attitudes towards mathematics, we also planned to analyze our results by gender, ethnicity, and socioeconomic status (Bolger & Kellaghan, 1990; National Center for Education and Statistics, 1994; Nettles & Nettles, 1995; O'Neil & Brown, in press).

Methods

Subjects

Our study design was based on a larger pilot study conducted by the state of California and was largely dependent on their volunteer sample. Our subsample consisted of 13 schools across the state and included 36 classrooms with more than 800 students. Within each school, three math classes were selected to represent the range of eighth-grade classes typically taught at that school. Forty percent of the students were White; the other 60% were evenly divided between African-American, Latino, and Asian-American ethnicities. Students were also evenly distributed by gender. In addition, students were distributed across SES levels, with about one-third of the students falling into each of low-, mid-, and high-SES-level categories. SES levels were defined using parental education as a measure of socioeconomic status. In reviewing our findings, it is important to note that participating schools were volunteers, selected from the larger state sample because they agreed to participate in further special pilot work.

Assessment Context

The CLAS assessment program featured a matrix sampling design and, at the middle grades (eighth-grade level), used a total of eight different forms for the mathematics assessment. Each form was composed of two distinct open-ended items and eight distinct multiple-choice items, yielding a total of 16 open-ended items and 84 multiple-choice items for estimating school-level performance. For the general administration of CLAS, students were randomly assigned to one of the test forms, and test forms were equally distributed in schools and classrooms. The goals of the assessment thus were not to provide individual student-level results but rather to estimate school-level performance. In addition, students in the pilot schools also took a pilot-level common form assessment, one of the eight forms from the general administration and the same for all students.

As mentioned, each assessment form consisted of two sections, the first containing two open-ended tasks and the second composed of eight multiple-choice items. The two open-ended tasks were designed to pose authentic, relevant problem situations for students to solve; the multiple-choice items were intended to assess mathematical thinking. Although we are unable to provide the actual assessment items (for confidentiality reasons), Figures 1 and 2 present sample open-ended and multiple-choice problems taken from the CLAS Addendum (a document disseminated by the California Department of Education which includes sample CLAS-type exercises—California State Department of Education, 1992). Since the math CLAS contained both novel open-ended and more familiar multiple-choice problems, it offered an excellent opportunity to compare student attitudes towards these two different types of tasks and to examine students' understandings of the performance criteria for each question type. Clearly, how students approach a given task will be influenced by their attitudes towards this task and their expectations for what is required of them in this task.

> Last year **Eat It Up Burgers** employed 5 workers for 5 hours a day. They claimed they served 4 million burgers last year. Is this a reasonable claim? Explain your answer.

Figure 1. Sample CLAS open-ended item.

> Study this sequence of numbers:
>
> 1, 3, 4, 7, 8, 9, 13, 14, 15, 16, 21, 22, 23, 24, __, __
>
> Which one of these could be placed in the spaces to continue the pattern?
>
> A. 25, 26
>
> B. 25, 31
>
> C. 25, 30
>
> D. 31, 32

Figure 2. Sample CLAS multiple-choice item.

Instrumentation and Procedure

The data presented in this article were collected using both student surveys and individual student retrospective interviews. All students in sampled classrooms completed a survey on the day following the administration of the pilot-level CLAS. The survey solicited students' views on a number of issues, including their attitudes towards open-ended as compared to multiple-choice tasks on the CLAS. In addition, in-depth student interviews were conducted with six students randomly chosen from each classroom. The individual student interviews allowed researchers to obtain more detailed information on student responses to the open-ended and multiple-choice tasks included in the assessment. Think-aloud protocols asked students to recreate their thinking processes and expectations as they approached and tried to solve one of the two open-ended tasks and the first multiple-choice item included on the pilot-level CLAS. The interviewers also asked students to explain how they thought each task would be scored and their relative preferences—along a number of affective dimensions—for open-ended versus multiple-choice problems.

Results

Data Coding and Reliabilities

Categories for coding responses to open-ended questions on our instruments were derived by reviewing a sample of responses; the major themes and/or key ideas so-identified were then operationally defined and used to categorize each response.

Interrater reliability was established by double-coding a set proportion of responses, with the proportion varying depending on the complexity of the coding categories. Because of their complexity, all student retrospective interview responses dealing with how students approached individual assessment tasks were coded by two raters; interrater agreement on these questions ranged from 0.69 to 1.0, with a median of 0.91. Only a sample (25%) of the student interview open-ended attitude responses was double-coded since coding categories for these responses were more straightforward. Interrater agreements on these items confirmed this judgment, as rater agreements ranged from 0.85 to 1.0, with a median of 0.94.

Data Analyses

Data for this article were coded and the analyses conducted at the individual student level. Descriptive statistics were calculated to characterize both students' attitudes and students' approaches to the CLAS tasks. These analyses highlight many differences in student attitudes between the two different problem types as well as differences in student perceptions of the performance criteria necessary for each problem type. In addition, analyses comparing students by school type (suburban, urban, and rural) and by gender were conducted. The school-type variable was used as a proxy for socioeconomic status, because of problems with missing individual-level data on students' socioeconomic status: The suburban schools were in high-wealth communities; the urban schools were in economically disadvantaged, inner-city communities; and the rural schools were of mixed socioeconomic status. These subgroup analyses did not yield significant results, except as indicated below. Analyses by ethnicity could be not conducted because high levels of missing ethnicity data in predominantly African-American schools would have biased the analyses.

Data Results

Expressing interest and challenge. Slightly more than half of the students who completed the survey—and almost two-thirds of those expressing a preference—indicated that open-ended questions are more interesting to solve than multiple-choice questions. (See Table 1 for a summary of survey results and Table 2 for a summary of interview results.) In addition, more than half the students surveyed reported trying harder on open-ended problems than on multiple-choice problems, while only 11% reported trying harder on multiple-choice problems. Furthermore, students who participated in the retrospective interviews reported overwhelmingly (83%) that open-ended questions are more challenging to answer than multiple-choice questions. Asked why, nearly half of these students (49%) mentioned that open-ended questions are more challenging because they cause students to think harder, are more difficult, or are more complicated to answer. In addition, approximately one-third (37%) of the students pointed out that open-ended questions require them to explain their answers by showing their work or by communicating their mathematics knowledge verbally. About one-fourth of the students also stated that open-ended questions are more challenging

Table 1

Student Survey Results: Percentage of Students Reporting Various Attitudes Comparing Open-Ended Versus Multiple-Choice Questions

Survey question	Multiple-choice	About the same	Open-ended
Which type of question did you find most *interesting*?[a]	29	19	51
On which type of question did you *try* harder?[b]	11	34	55
Which type of question do you think *showed better* what you know about math?[c]	40	24	36
On which type of question do you think you *did* better?[d]	68	17	14
Which type of question did you *like* better?[e]	60	22	17

Note. Only single responses allowed. Total number of students surveyed was 792.
[a] n = 773. [b] n = 777. [c] n = 776. [d] n = 774. [e] n = 774.

Table 2

Student Interview Results: Percentage of Students Reporting Various Attitudes Comparing Open-Ended Versus Multiple-Choice Questions

Interview question	Multiple choice	No difference	Open-ended
Which kind of question makes you think harder or is more challenging?[a]	8	9	83
Which kind of question best lets you show what you know about math?[b]	34	11	55
Which kind of question was easier for you to understand what to do?[c]	65	14	21
Which kind of question do you like better?[d]	54	14	32

Note. Only single responses allowed. Total number of students interviewed was 212.
[a] n = 210. [b] n = 208. [c] n = 209. [d] n = 211.

because students have to create an answer on their own (22%) or because specific answers are not provided (24%). For example, one interviewed student reported, "Open-ended is more challenging because you have to put your mind to it and think about it real hard, and then you have to show different types of things to make it easy to understand the problem: graphs, and charts, and stuff like that."

Asked what they liked best about open-ended items, interviewed students reported that they like being able to create their own answers (14%), having an opportunity to explain their answers (27%), and using graphs or diagrams in their responses (12%). Students also reported liking open-ended items because they are challenging (29%). One student remarked, "I like the thinking and the showing. I also like the creativity that you have to put into it [to show] the solutions to the question."

Claims that state that open-ended problems are more motivating to students than traditional problems are bolstered by these results. In addition to finding open-ended tasks more interesting, students' perceptions of the relative challenge in open-ended versus multiple-choice tasks mirror the intentions of proponents and developers of alternative assessments. Open-ended items apparently require students to actively accomplish complex tasks. These findings are also encouraging since alternative assessments strive to tap higher order cognitive processes or problem-solving skills and to encourage students to create or produce a response—something of which students are apparently aware.

Showing math capability. Unlike the previous results—in which students clearly cited open-ended tasks as more interesting and challenging—questions comparing how well the two types of tasks show students' knowledge and abilities garnered mixed results. Student survey results found students almost evenly divided regarding which type of question shows better what they know about math. About 40% of the students reported that multiple-choice tasks better let them show what they know, 36% chose open-ended tasks, and 24% of the students reported no difference between the two types of tasks. However, interviewed students (who reviewed their responses to the multiple-choice and open-ended tasks prior to responding to the question) were more likely to state that open-ended tasks best enable them to show what they know. More than half of the students (55%) so indicated. "Open-ended [best lets you show what you know] because you have to use graphs and explain your answer," observed one student. Almost two-thirds (65%) of these students stated that open-ended questions best show what they know because these questions allow them to explain their answers. Additionally, students reported that open-ended questions indicate math knowledge because these questions are more challenging (20%), they do not include answer choices (12%), and students must create their own answers (8%).

Inconsistent findings between survey and interview responses may simply signal that students are unsure which type of question best lets them show what they know because they are as of yet unfamiliar with how understanding in open-ended problems is assessed. In contrast, the manner in which multiple-choice solutions are scored is clear and straightforward. Further, interviewed students who were able to review the tasks and explain their solutions to the interviewers may have favored open-ended tasks over multiple-choice tasks in this question because during their interview they had just "shown" (the interviewer) what they knew.

Understanding what's being asked. Since students frequently felt that open-ended questions were more challenging, it is not surprising to find that almost two-thirds of the students interviewed (65%) felt it easier to understand what to do in *multiple-choice* rather than open-ended questions. "You know one of these is going to be the right answer so it's easier to mark your answer," explained one student. Of the students reporting that multiple-choice questions make it easier to understand what to do, many attribute this ease to the answer choices being given (39%), the actual questions being easier (25%) or stated more clearly (18%), or the lack of a requirement to explain their answers (12%).

Students' responses, in fact, are similar to some of the criticisms mounted against multiple-choice questions in the past (e.g., neglect of complex thinking and problem solving). The familiarity of the multiple-choice format—in contrast to the relative newness of open-ended tasks—may also play a role in students' beliefs that these types of questions make it easier to understand what to do. Many students did in fact express their familiarity with multiple-choice tasks, stating such things as, "multiple-choice [is easier to understand what to do] because we do stuff like multiple-choice every day."

Performing better. Consistent with their opinions about which type of problem is easier to understand, more surveyed students felt they perform better on multiple-choice questions (68%), compared to those reporting they perform better on open-ended questions (14%). The familiarity of the multiple-choice format may again account for this judgment, since students understand well how multiple-choice tasks are scored but are less familiar with open-ended tasks and the processes they entail. In addition, since students find open-ended tasks more challenging, they therefore—by contrast—find multiple-choice tasks easier.

Liking problems. While the majority of students surveyed (61%) reported they are not frustrated by problems with more than one answer, only 39% of the students surveyed stated that they like problems with no obvious solution. In addition, only 17% agreed that they like problems that take a lot of time to solve. In particular, males were more likely than females to report liking problems with no obvious solutions, $F(1, 599) = 9.22$, $p = .003$. Also, suburban students were less likely than other students to report frustration when faced with problems having more than one answer, $F(2, 789) = 4.37$, $p = .01$.

Confirming these responses, survey results indicate that students like multiple-choice questions (60%) better than open-ended questions (17%). Similar patterns were found in the interview results: Just more than half of the students interviewed (54%) stated they like multiple-choice tasks better than open-ended tasks (versus 32% who like open-ended tasks better). The reasons students gave for liking multiple-choice questions better included that these questions are easier (58%) and that the answer choices are given (41%). Reasons given for disliking open-ended questions included the difficulty of the questions (42%), the need to explain the answers (19%), problems in understanding the questions (10%), and the lack of question clarity or the lack of information given in open-ended questions (8%). Students' aversion to open-ended problems can be summarized well by one student's statement: "They're hard."

Thus, although students report open-ended items to be more interesting and challenging, they still prefer multiple-choice items. Students' beliefs that multiple-choice problems are easier stem partly from the availability of response choices for these types of problems; in addition, multiple-choice problems may be perceived as easier because of students' familiarity with them as compared to the relative newness of open-ended items.

Perceiving performance criteria. In the retrospective interviews, students were asked what they thought their teachers would be looking for as they scored student responses to open-ended problems: How would the responses be graded? Parallel questions were posed concerning the multiple-choice problems.

Students' answers to the interview questions were coded to indicate whether or not students mentioned any of the following dimensions as important in grading: (a) the correct answer, (b) the steps students used to solve the problem, (c) students' use of graphs, charts, or diagrams, and (d) the depth of students' explanations and understanding.

Table 3 shows the distribution of students who mentioned each of the four criteria by task type. For novel open-ended items, almost half (46%) of the students mentioned the importance of the quality or depth of their explanations; 51% mentioned attention to their use of diagrams, graphs, and other visuals; 26% mentioned that the steps of their solutions would be important; and 26% thought the correct answer was an important element in scoring. For example, one student said she would expect her open-ended task to be graded by "how I did the graph, or how well I tried to explain myself." In contrast, when asked about the multiple-choice items, 45% of the students indicated that scorers would be looking for the correct answer, while 34% mentioned the steps used to solve the problem (perhaps recognizing that if their method was not correct, they were unlikely to get the right answer). Diagrams were mentioned by only 1% of the students, and the importance of explanation was mentioned by only 16% of the students. Common responses to questions regarding what teachers would be looking for on multiple-choice items included comments such as "just if I got it right" or "the right answer."

Table 3

Student Interview Results: Percentage of Students Mentioning Each of the Four Grading Dimensions for Open-Ended and Multiple-Choice Tasks (N = 192)

Grading dimensions	Multiple-choice	Open-ended
Correct	45	26
Steps	34	26
Diagrams	1	51
Explanation	16	46

Note. Multiple responses allowed. Total number of students interviewed was 212.

In addition to the information coded from the student interviews, students also were asked directly whether they think or do anything differently when responding to open-ended problems as compared to multiple-choice problems. Two-thirds of the students (67%) reported approaching open-ended tasks differently. When asked *why* they reported differences, 40% of the students replied that they have to explain their thinking and 37% reported that they have to think harder on open-ended items. Other answers included the need to create their own answers (11%), the use of diagrams (11%), and the lack of given responses from which to choose (13%).

It appears then that students do perceive different expectations for their responses on open-ended versus multiple-choice problems. Multiple-choice items are associated with the use of appropriate algorithms and the determination of a correct answer; open-ended items, on the other hand, are allied with the use of diagrams and the need to explain one's results. Students apparently are aware that open-ended tasks require a different type of approach than do the more familiar multiple-choice tasks.

Reasoning. Another aspect of students' approaches to CLAS tasks was shown in the lines of reasoning students used when solving the tasks. Do students pursue a mathematics-based reasoning approach to solve a given problem or do they use a non-mathematically-oriented trial-and-error or guessing approach? Are there differences in how students approach novel open-ended tasks and how they approach familiar multiple-choice tasks?

Student retrospective interview responses were coded for the type of reasoning students used. "Mathematics-based reasoning" was defined as that which utilized disciplinary concepts (rightly or wrongly) or strategic lines of reasoning based on mathematical thinking. For example, one student reasoned about a problem: "Well, the area of the larger square minus the smaller square should give you the shaded area and to get area from perimeter you. . . ." In contrast, random "trial-and error" approaches were defined by their lack of logic from a mathematical perspective. Many students attempted to take the numbers from a problem and play with them in order to come up with one of the given multiple-choice alternatives. For example:

> First I was thinking about how I was going to find the area . . . I didn't really know how to do this. But they told me the perimeter of the larger square and they told me the perimeter of the smaller square. I was kind of thinking to do something like divide, I wasn't sure. I didn't think if I multiplied twenty and sixteen, that would come out to one, two, four or nine.

These latter types of trial-and-error responses were associated with students who admitted guessing (e.g., "I looked at the answers, and I thought four or nine would be all right, but I wasn't sure, so I picked four feet."). These results are displayed in Table 4. In the open-ended problems, students overwhelmingly followed some mathematics-based reasoning approach (whether correct or incorrect) rather than using a trial-and-error or guessing approach: Only 4% of student responses were coded as guesses. In contrast, 35% of the students used a trial-and-error or guessing approach on the multiple-choice items. Suburban students were more likely than other students to use a correct, mathematics-based reasoning approach on both open-ended problems, $\chi^2(4, N = 139) = 14.2, p = .007$, and multiple-choice problems, $\chi^2(4, N = 187) = 44.2, p < .0001$.

Table 4

Student Interview Results: Percentage of Students Using Guessing or Mathematics-Based Reasoning Approach for Open-Ended and Multiple-Choice Tasks

Task approach	Multiple-choice[a]	Open-ended[b]
Trial-and-error or guessing	35	4
Mathematics-based reasoning	65	96

Note. Only single responses allowed. Total number of students interviewed was 212.
[a] $n = 187$. [b] $n = 139$.

Students also were asked in the interviews whether they guess on tests; 82% of the students reported doing so. When asked on which they guess more—open-ended or multiple-choice problems—students reported guessing more on multiple-choice problems by a margin of 83% (multiple-choice) to 13% (open-ended). When asked *why* they guess more on the indicated problem, 61% of the students reported that they guess more on multiple-choice problems because these problems "give you a selection of answers."

Coding of students' comments during the retrospective interviews indicated that 35% of all students mentioned using the multiple-choice response alternatives as prompts to help them solve the problems. For instance, one student explained, "When you finish everything [in a multiple-choice question], you have some back-up possible answers; [with] open-ended you don't know if it's really right because you don't have anything to back it up." Students reported using the given alternatives to prompt their responses in a variety of ways: Some students described using the alternatives to check their work; others indicated selecting the "closest" alternative to the answer they had computed; and still others reported guessing from the given alternatives. A number of students also indicated they used the possible alternatives as a starting point, working backwards from these alternatives to the initial problem.

Discussion and Conclusion

From the data presented here, we can begin to paint a picture of how students felt about and approached the mathematics tasks found on the California Learning Assessment System. As indicated in the introduction, the literature surrounding alternative assessment suggests that students will find alternative assessments more meaningful and more motivating than traditional multiple-choice tasks, and furthermore, that students who engage in authentic tasks are likely to be more motivated to learn in school. In general, students' perceptions of open-ended items are consistent with major aims of proponents and designers of alternative assessment. Furthermore, our analyses show little difference in perceptions by gender or by our limited indicator of socioeconomic status, although those differences that emerge suggest that economically advantaged students are more comfortable with the demands of these new forms of assessment, both in their attitudes towards problems with more than one solution and in their abilities to use appropriate mathematical thinking to solve problems. This is not surprising in that a number of researchers have suggested children of poverty are far more likely than their advantaged peers to have been subjected to the "drill and kill" rote curriculum spawned by overemphasis on standardized multiple-choice tests (Darling-Hammond, 1995; Herman & Golan, 1991).

Study results show that students find alternative assessment items more interesting and challenging than multiple-choice items; students try harder on these items; and they recognize that open-ended items require them to think harder, explain their thinking, and communicate their understanding of mathematical knowledge. At the same time, however, students do not necessarily like such challenges. In fact, students express a preference for multiple-choice items. They find multiple-choice items easier to understand and believe that they perform better on such items. These preferences may in part be due to the relative newness of open-ended items as compared to the comforting familiarity of multiple-choice items, as well as possibly due to a lack of student understanding of how their performance on open-ended items will be assessed.

We can also report that alternative assessment, as represented by the open-ended tasks included on the CLAS assessment, is achieving at least some of its aims. Although it was unclear to students which type of task best illustrates their mathematical knowledge, students did seem to understand the differences in approach necessitated by open-ended versus multiple-choice problems. They know that open-ended problems emphasize students' use of explanatory materials (e.g., graphs, charts, diagrams, and the quality of their explanations) and focus less on algorithms and the correct answer. Students seem to follow a mathematics-based reasoning approach—correct or otherwise—in responding to open-ended items; they approach multiple-choice items both logically and by using a trial-and-error or guessing approach. Furthermore, by not giving students answer alternatives, open-ended items clearly inhibit simple guessing strategies. Students cannot use answer alternatives to prompt their responses, marking different approaches to solving multiple-choice tasks and more novel open-ended tasks.

While values and student perceptions may change as open-ended tasks become more familiar to students with practice and exposure, results here suggest that alternative assessment, in the eyes of students, is living up to many of its claims.

References

Baker, E. L., Freeman, M., & Clayton, S. (1991). Cognitively sensitive assessment of subject matter: Understanding the marriage of psychological theory and educational policy in achievement testing. In M. C. Wittrock & E. L. Baker (Eds.), *Testing and cognition* (pp. 135-153). New York: Prentice-Hall.

Bolger, N., & Kellaghan, T. (1990). Method of measurement and gender differences in scholastic achievement. *Journal of Educational Measurement, 27,* 165-174.

California State Department of Education. (1992). *Addendum to a sampler of mathematics assessment*. Sacramento, CA: Author.

Carroll, J. B. (1975). *The teaching of French as a foreign language in eight countries*. New York: John Wiley.

Darling-Hammond, L. (1995). Equity issues in performance based assessment. In M. T. Nettles & A. L. Nettles (Eds.), *Equity and excellence in educational test and assessment* (pp. 89-114). Boston, MA: Kluwer.

Herman, J. L., & Dorr-Bremme, D. (1983). Uses of testing in the schools: A national profile. In W. Hathaway (Ed.), *Testing in the schools. New directions for testing and measurement* (No. 19, pp. 7-17). San Francisco, CA: Jossey-Bass.

Herman, J. L., & Golan, S. (1991). *Effects of standardized testing on teachers and learning—Another look* (CSE Tech. Rep. No. 334). Los Angeles: University of California, Center for Research on Evaluation, Standards, and Student Testing.

Herman, J. L., Klein, D. C. D., Heath, T. M., & Wakai, S. T. (1995, April). *Assessing equity in alternative assessment: An illustration of opportunity-to-learn issues*. Presentation at the annual meeting of the American Educational Research Association, San Francisco.

Herman, J. L., Osmundson, E., & Pascal, J. (1996). *Evaluation of the Los Alamos National Laboratory Critical Issues Forum*. Los Angeles: University of California, Center for the Study of Evaluation.

Kellaghan, T., & Madaus, G. F. (1991). National testing: Lessons for America from Europe. *Educational Leadership*, *49*(3), 87-93.

Lomask, M., Baron, J., Greig, J., & Harrison, C. (1992, March). *ConnMap: Connecticut's use of concept mapping to assess the structure of students' knowledge of science*. Symposium presented at the annual meeting of the National Association of Research in Science Teaching, Cambridge.

Madaus, G. F. (1991). The effects of important tests on students: Implications for a national examination system. *Phi Delta Kappan*, *73*, 226-231.

National Center for Education Statistics, U. S. Department of Education. (July, 1994). *The NAEP 1992 technical report* (Rep. No. 23-TR20). Washington, DC: Author.

Nettles, M. T., & Nettles, A. L. (Eds.). (1995). *Equity and excellence in educational test and assessment*. Boston, MA: Kluwer.

O'Neil, H. F., Jr., & Brown, R. S. (in press). Differential effects of question formats in math assessment on metacognition and affect. *Applied Measurement in Education*.

Shepard, L. (1991). Will national tests improve student learning? *Phi Delta Kappan*, *73*, 232-238.

Smith, M. L., & Rottenberg, C. (1991). Unintended consequences of external testing in elementary schools. *Educational Measurement: Issues and Practice*, *10*, 7-11.