ALTERNATIVE ASSESSMENT: STATE ACTIVITY, INTEREST, AND CONCERNS

CSE Technical Report 322

Pamela R. Aschbacher

UCLA Center for Research on Evaluation,
Standards, and Student Testing

May 1991

The research reported herein was conducted with partial support from the U.S. Department of Education, Office of Educational Research and Improvement, pursuant to Grant No. G0086-003. However, the opinions expressed do not necessarily reflect the position or policy of this agency and no official endorsement by this agency should be inferred.

Please address inquiries to: CSE Dissemination Office, UCLA Graduate School of Education, 405 Hilgard Avenue, Los Angeles, California, 90024-1521

Why Are States Interested in Alternative Assessment?

Billions of dollars are spent each year on education, yet there is widespread dissatisfaction with our educational system among educators, parents, policymakers, and the business community. Efforts to reform and restructure schools, stemming from *A Nation at Risk* (National Commission for Educational Excellence, 1983) and other reports critical of the quality of American education, have focused attention on achievement outcomes and the role of assessment in school improvement. Many educators have strongly criticized the measures used to monitor student performance and evaluate programs for failing to assess significant learning outcomes and for thereby undermining curriculum, instruction, and policy decisions (Baker, Freeman, & Clayton, 1990).

In the past, multiple-choice measures have been relied upon for their ease of administration and scoring as well as the extensive test theory and statistical knowledge base that supports their use. But critics have noted that most commonly used tests, whether typical standardized norm-referenced tests, objective-based tests, or many of the tests created by teachers to grade students, tend to emphasize mere recall of knowledge and provide little information about the level of student understanding or quality of thinking (Nickerson, 1989). It is argued that without a clear window on students' complex thinking and problem-solving skills, not only do we fail to evaluate our students and instructional programs adequately, but we also communicate to teachers, parents, and students that such untested skills are not very important (Baker, Freeman, & Clayton, 1990). Thus, such tests can have a deleterious influence on the classroom—on how information is presented, learned, and retained. These tests assess learning in an artificial, decontextualized manner that is removed from the ways students actually learn and will need to apply knowledge outside the classroom (Resnick & Klopfer, 1989). Since the influence of testing on curriculum and instruction is now widely recognized, policymakers, large-scale testing directors, curriculum experts, teachers, and others are turning to alternative assessment methods as a tool for school improvement.

What are the Desired Characteristics of Alternative Assessment?

There is no single method of alternative assessment. Rather, there are many different approaches, such as portfolios of student work over time, exhibits or displays of knowledge and skills, open-ended questions with no single right answer, and hands-on experimentation or demonstration. The common threads uniting these approaches as alternative or performance-based assessment are:

- students perform, create, produce or do something that requires higher level thinking or problem-solving skills (not just one right answer);
- assessment tasks are also meaningful, challenging, engaging *instructional* activities;
- tasks are set in a real-world context or close simulation;
- process and conative behavior are often assessed as well as product; and
- the criteria and standards for performance are public and known in advance (Aschbacher & Herman, 1990).

The concept of performance-based assessment has seductive face validity. It promises to exemplify the broader and deeper range of desired performance more than multiple-choice questions ever could and to inform change towards better instruction and evaluation. But with so much at stake (time, money, widespread

training, curriculum changes, public credibility, and students' futures, to name a few), many educators and testing experts are cautious. They are concerned whether these new directions are better than previous ones and will not fall prey to the same kinds of problems that have plagued traditional multiple-choice testing when it is made to serve accountability and instructional purposes, discussed by Linn, Graue, and Sanders (1990).

How Can States Find Out More about Current Implementations of Alternative Assessment?

An important step for those who would consider or develop new assessment methods is to examine what others have already tried to accomplish, what barriers or problems they have encountered, what success they have had, and what issues or concerns remain to be addressed. Unfortunately, many state and district testing offices are suffering serious budget constraints and do not have the staff to search the literature or systematically monitor new efforts across the country, nor do they have the money to attend many conferences and meetings to obtain information and technical assistance. The purpose of this paper is to share information about current state interest, activity, and concerns related to performance assessment.

Which States are Involved in Alternative Assessment and What are They Doing?

We conducted telephone interviews with the directors of student assessment programs (or policy and planning) in each of the 50 states during spring 1990. Each state testing officer was interviewed about the state's activity or interest in alternatives to multiple-choice testing at the state level, defining "alternative assessment" to include open-ended written questions, hands-on experiments, performances or exhibits, portfolios of work, and so forth. We also asked about the major issues, concerns, or barriers they faced or anticipated. Since the focus was on *new* approaches, simple direct writing assessments, which have been around for years and are currently used in over half of the states, were not the focus of this telephone survey unless they involved something innovative, such as writing in the content areas, portfolios of writing, or extended writing and editing over time. A separate survey of states was conducted later to obtain information about direct writing assessments (see Appendix A and Table 2). Table 1 summarizes current alternative assessment activity or interest by grouping the 50 states according to their stage of involvement. To facilitate networking, the table notes content areas of interest.

As of the end of 1990, about half of the states are involved to varying degrees in alternative assessment and are rather evenly divided among three categories:

1. Those who have had some alternative assessment in place for several years (although some of these efforts are quite modest in scope, such as North Carolina's few geometry proofs, and some are not used every year);

2. Those who are actively developing and/or piloting alternative assessments; and

3. Those who are exploring possible alternative assessment ideas among state assessment staff or state level committees, or who are supporting districts that are trying to develop new local assessment methods.

As Table 1 illustrates, most of the assessment activity and interest are in the areas of math (21 states) and science (17), followed by social studies (14), expanded

writing (11), and expanded language (8). A few states also are involved with or interested in a performance-based approach to assessing reading, music, art, physical education, and employability skills. The direct writing assessment survey revealed that at least three states are considering assessing writing and other skills in cooperative group settings.

Despite apparent interest among many of their testing directors, about half the states are not currently implementing nor planning to implement alternative state assessment programs within the next several years. Their reasons for inaction are varied: some do not feel alternative assessment is appropriate or feasible at the state level, some are waiting to see what will prove feasible in other states before committing their limited resources to new approaches, others are awaiting a change of governor or state education officer, and many have no staff or budget to pursue alternative assessment possibilities regardless of their interest.

What are the Major Concerns That States Have about Alternative Assessment

Four major areas of concern were noted by states in discussing alternative assessment: costs, logistics, technical concerns, and lack of support for innovation. Each of these is discussed separately below.

Cost

As might be expected given the current poor economic climate, the most frequently cited obstacle in most states was cost. Testing directors noted that the costs of alternative measures can be enormously higher than standardized multiple-choice tests (for example, some states mentioned figures of $3.00 to $10.00 per student for performance measures versus $1.50 for some typical standardized tests). These higher costs reflect increases at all stages of development and administration:

- training the State Education Agencies staff,
- developing the actual tasks and scoring criteria,
- public relations,
- training people to administer the tests and to score them,
- training teachers and administrators regarding new approaches to assessment and appropriate curriculum revisions,
- providing equipment (such as hands-on science kits) and transporting it to school sites,
- administering the measures,
- rating or scoring the products, processes, and/or affect to be assessed,
- providing substitutes for teachers involved in training, administering, and/or scoring,
- processing and analyzing complicated data, and
- reporting and explaining the results.

Some of these costs are essentially one-time only costs associated with new programs; others reflect the labor intensity of alternative assessment approaches.

For example, some testing officers doubt that teachers are ready to administer hands-on science tests yet since so few of them currently use such approaches in their own classrooms. They felt that one motivational workshop would not be sufficient. Teachers need concrete, intensive training, which may be quite costly and time-consuming for the state to develop as well as to deliver. Doing so also requires some performance-based assessment expertise at the state education agency, which may take some time to attain.

A number of solutions were suggested by testing directors to overcome the problem of high costs. Among these is spreading the cost over other budgets, such as staff development and curriculum development. Several states noted that implementing an alternative assessment program had required much teacher training, but that its impact on curriculum and instruction had been worth it. Not only had teachers been trained in new assessment approaches that they could utilize in instruction, but also their morale and incentive to take risks had been raised significantly. Testimonials aside, it is important to consider whether the specific teacher training required for a given alternative assessment comprises (or could be revised to comprise) the type of staff development truly needed in a particular district or state.

Another solution to the cost barrier is to test only a sample of the school population rather than every pupil. This approach is far more feasible logistically as well; however, some policymakers tend to distrust sampling and to insist on testing every pupil. In other cases, sampling is eschewed because diagnostic information on individuals is desired and there are only sufficient resources for one type of test.

Alternative assessment methods are labor-intensive, and scoring can be a particularly costly aspect. States differ in their approaches to this problem. Some feel it is cheaper to utilize professional scoring services, while others feel it pays to train their own teachers, particularly when one objective is to educate teachers and to encourage certain types of instructional practices (e.g., writing in the content areas).

On the other hand, using the state's teachers as test administrators and raters is a costly commitment. Beyond teacher time and potential substitute teacher pay, the cost of travel to regional training and/or scoring sites can be quite high, causing states to weigh the technical and fiscal costs and benefits of regional versus residential scoring. Remote scoring, with scorers networking through computers and receiving immediate reliability feedback, may eventually help to lower costs, but development of viable procedures remains to be accomplished.

Some states mentioned that they are able to keep costs down by maintaining excellent working relations with local education agencies, who may be willing and able to share the costs as well as the benefits. However, in many states the school districts are also suffering from a poor economy and may well resent mandated testing programs that require them to take staff time away from other important tasks and to pay for substitutes, travel and training.

Utilizing volunteers from colleges, community organizations, and the private sector has helped a couple of states keep their costs down and involved the larger community in education issues. Unfortunately, this approach requires particular effort to maintain sufficiently standard test administration practices and reliable scoring.

A few districts have been able to create models of innovative assessment and instruction at the local level with funds from private foundations or corporations. Businesses and funding agencies could provide a more effective solution to the cost barrier at the state level by supporting innovative low-cost models that address state

and local assessment issues and by providing extensive visibility for successful projects.

Logistics

Many of the states that do not currently have an alternative assessment program cited the complicated logistics of such testing as a significant barrier. "Too much paper to shuffle," "too much equipment to keep track of and move from site to site" and "too much class time spent on testing" are common refrains.

Sampling or voluntary inclusion of districts or schools are approaches that have reduced the logistical problems for some states; however, sampling is not a viable solution for some. A number of states, such as Illinois and New Jersey, have a legislative mandate to test all students at given grade levels. Some states prefer testing every pupil for accountability purposes, others want to promote individual diagnosis, and some feel that the results are taken more seriously when every school is involved in the testing. Many policymakers are interested in school reform at the building and classroom level, so they reject sampling plans in favor of census testing yet mandate quick turnaround of results to schools and districts. Such an approach greatly increases the costs and logistical problems associated with performance-based assessment. For example, one state this year will be spending a million dollars to score 300,000 essays and report the results within 60 days.

Some states are experimenting with testing students in pairs or small groups to reduce the amount of equipment, paper, and time involved as well as to examine group interaction skills. Others rotate the content areas tested across years. Some are considering testing integrated subject areas (for example, writing in social studies, or math combined with science) to conserve time and simulate real life problem-solving. This approach, however, is only in the early stages of development (see Baker, Freeman, & Clayton, 1990, and Aschbacher, 1991, for discussion of an approach to writing in social studies that may eventually prove useful in this regard).

The logistics of transporting, storing, and setting up the special equipment needed for some alternative measures can be a problem. For example, hands-on science kits frequently contain parts that may become lost or broken and need to be replaced. Many schools could also have difficulty finding room in which to set up testing materials not administered in regular classrooms. Hawaii has a novel approach to transporting the extensive equipment that comprises its performance-based graduation exam given to students who have not passed the standard exam: a traveling van visits each school where it is needed.

The complicated schedule of classes in many middle schools and high schools can also create a logistical problem unless tests are carefully designed to fit the length of regular class periods, or the structure of the testing days is changed.

Another logistical problem is assuring the timeliness of results. If new methods are to have a strong influence on classroom practices, teachers and administrators must receive the results in a timely fashion. By most accounts, it takes longer to get results scored, analyzed, and reported with alternative assessment than with the standardized multiple-choice exams that take advantage of highly automated scoring systems. Particularly when classroom teachers are used as raters, obtaining quick feedback on a large population during the school year may be difficult. Where timeliness is particularly important, states or districts may find it necessary to compromise by reducing the complexity of the scoring schemes (see Baker, Aschbacher, Niemi, and Chang, 1991, regarding trade-offs in complexity of scoring schemes) or the analyses of results or relying on commercial scoring services. As experience with alternative assessment grows, scoring and reporting probably will become more efficient over time.

Technical Concerns

Many states are aware that technical concerns such as reliability and validity are not just of interest to statisticians but are *central* to their decision whether to include performance-based assessment in their statewide assessment programs. One state testing officer says:

> *It's important that tests have credibility so that we don't waste time arguing over data, but rather schools can take action to improve.*

In today's economic climate it is clear that no state has the time or money to waste on assessment efforts that have low credibility and little likelihood of providing a basis for school improvement.

Whether the test is an alternative assessment or traditional multiple-choice, reliability of administration and scoring and validity of measures are especially important when there are significant consequences tied to students' "competence" on the test (i.e., in "high stakes" testing situations). States now find it necessary to document that they have used the most current and well-respected approaches to establishing reliability and validity in case they must defend their methods in court.

Some state testing officers noted that judging performance, process, or conative behavior is not as simple, straightforward, or efficient as some would claim and may require much consideration before judges attain consensus on standards and criteria. Sound procedures require, among other things, multiple raters, highly structured rating systems and rater training, content validation by subject matter experts, multiple opportunities for students to pass, equating, and security of test content.

Generalizability of performance is another concern raised by alternative assessment. How valid is it to generalize from a student's performance on one or a few tasks to the complete domain of interest? Good multiple-choice tests typically contain at least five to eight items per objective carefully selected from a well-specified and usually fairly narrow domain in order to solve this problem. Performance assessments, however, are usually comprised of fewer but larger and more integrated tasks, leanly sampled from the target domain. Research is clearly needed on these issues.

A few testing directors noted that alternative assessments could potentially result in narrowing of curriculum much as multiple-choice tests have done. Even with alternative assessments, it is important to avoid a situation in which high stakes encourage teachers to focus too heavily on specific tasks known to be on the test at the expense of other important content or processes, resulting in students superficially or only temporarily learning information with little deep understanding or the ability to apply it. Assessment of the latter requires systematic samples from a broad range of desired content and process.

Many states also recognize that credibility and usefulness are directly linked to how the results are quantified and reported. Models of successful state level reporting of open-ended questions or portfolio results, for instance, need to be widely shared.

Support for Implementation

Support for the implementation of performance-based assessment is important in several quarters: the legislature and governor's office; the public,

parents and students; local teachers and administrators; and the state education agency.

Changes in the governor or state superintendent of schools affect what is proposed, implemented, and supported, and the anticipation of such changes tends to create a period of limbo while people wait for elections, decisions, or funding commitments. Few innovations may occur even after an election, since newly elected or appointed office holders may refrain from supporting new ideas such as alternative assessment unless there is strong popular support for them.

A few state legislatures are moving towards heavier oversight of schools through assessment and are now requiring comparisons of districts and schools where minimal statewide testing existed before. In some states there will be room for performance-based assessments of higher order thinking skills, but in general the focus in these states is on normative comparisons with traditional multiple-choice tests. Even the new NAEP tests will not be used because of past anomalies (e.g., with the reading test) or because of slow reporting.

State testing officers remarked that policymakers hold a wide range of views on alternative assessment. Some policymakers are suspicious that alternative assessment is simply an attempt by educators to "change the game" in order to dodge heavy criticism; some perceive traditional multiple-choice tests as more trustworthy and comprehensive than alternative measures; many are unfamiliar with new assessment methods; and some have been encouraged by the notion of measurement-driven instruction to push for rapid reform of assessment programs. Some state education agencies feel a need to compensate for policymakers' unbridled enthusiasm with a "go slow" approach.

According to a number of testing officers, the public and parents tend to be uninformed and somewhat uncaring about state testing unless it directly affects them, such as their child's graduation. They do need to feel assured that such tests are valid measures and will be scored reliably. Some fear that the more subjective scoring schemes and potential bias of human scorers for alternative assessments will have negative consequences for minority students. Others are concerned that minority students will perform more poorly on the new tests because of the lack of opportunities for instruction and practice in critical thinking skills; however, such consequences could argue for important changes in policy and instruction. Some parents also may be concerned that students tested with alternative assessments will not be adequately prepared for later gateway tests, such as the SAT, ACT or job entry tests.

One testing officer with experience in alternative assessments noted that some students, particularly twelfth graders, may feel imposed upon when asked to take performance-based assessments that they perceive as more difficult and time-consuming than the traditional multiple-choice tests, and they may not cooperate or be motivated to try on the tests. On the other hand, another state that has piloted some hands-on science lab experiments with middle school students found that they really enjoyed the experience and wanted to do more of it.

Some states in which teacher competency is a touchy political topic may feel reluctant to push ahead with new performance-based assessment to encourage new instructional strategies and content. In this milieu there may be negative consequences to using assessment techniques that are too far ahead of current instructional practices and that may threaten teachers with being viewed as "incompetent."

Securing acceptance of the logistics and outcomes of alternative assessment by teachers and administrators can be a problem. Despite their denigration of norm-

referenced tests as trivial, teachers and administrators often value the supposed objectivity of multiple-choice tests, and they tend to prefer the ease of administration associated with these traditional tests. For example, in one state which recently gave districts the opportunity to choose either the previously used norm-referenced tests or the new performance tests in reading and writing at the elementary grades, half the districts chose the old norm-referenced tests. This state's testing director concluded that a lot of training and promotion of alternative assessment may be necessary to achieve acceptance of these new methods.

The amount of time taken away from instruction, particularly at the high school level, may be a further disincentive for many teachers and administrators to participate in new performance-based assessment programs.

Where local education agency budgets allow, paying teachers for training and scoring during summer months can be an incentive to cooperate with a new testing program. Some states with alternative assessment programs have allowed the teachers who administer the tests to keep the hands-on science lab equipment for use in their own classrooms after the testing is finished, thus providing an incentive to participate and encouragement to improve instructional practices.

Support for performance-based assessment at the local level is clearly a function of budget problems at the state and local education agencies. As state education agencies try to contain their costs, they may pass some of them on to local education agencies, such as the cost of travel for teachers to attend training sessions or the cost of substitutes to cover teachers' classes while they are training for, administering, or scoring tests. However, many local education agencies are suffering budget cuts as well and are not in a position to be helpful. Reductions in local administrative jobs mean fewer people to do the same or greater work load, often with minimal support staff. In at least one state, the district director of research and evaluation has to oversee the administration of *seven* different testing programs and produce that many different annual reports of district testing results for accountability purposes.

One state noted that district ownership of alternative state assessment results increased when the state changed the nature of the sample of students tested. In the past, state level results based on a sample of schools had led teachers and administrators to think of performance-based activities as enrichment, not as mainstream assessment and instruction. When they switched to a sampling plan in which they could report performance data on individual school reports, schools were less likely to discount the results as irrelevant and the data provided the necessary leverage to impact instruction.

Many state testing directors feel that performance based assessment may have value to impact instructional practices in desirable directions, but they question whether such complicated and costly methods should be a part of state level assessment programs. What information is best collected by the state and what by local education agencies, and for what purposes? Some argue that alternative testing methods are best used only at the local level to inform instructional decision making. Others believe that state assessment policies and practices have such powerful influence on the implemented curriculum that they should include alternative assessments as a critical means to achieve school reform.

In states with existing large or complex testing programs, particularly where additional NAEP testing is scheduled in the near future, there is understandable reluctance to take on any additional testing programs at this time. Several states noted that any new testing now would have to be done without additional funding or staff time, thereby making it virtually impossible to accomplish.

Testing directors also noted that a move to alternative assessment as part of a state testing program would require expansion of expertise in the state department of educations and, perhaps, unknown changes in the relationship of the assessment division to the division of curriculum and instruction.

Summary

A survey of state testing directors revealed strong arguments in favor of alternative assessments as a potent tool to improve curriculum and instruction and thereby raise student outcomes. However, a number of obstacles to state level use of performance-based assessments were cited: cost, logistics, technical concerns, and poor support for implementation. Despite these concerns, about half the states are currently conducting, developing, piloting, or examining alternative assessments. Ten of the states that are currently uninvolved indicated specific interests, such as math portfolios, which they might like to explore in the future. State testing directors offered a few solutions to the problems cited above and noted the need for continued research on technical and practical issues as well as a need for greater collaboration, documentation and sharing of successful efforts around the country.

References

Aschbacher, P.R. (1991, April). *The effects of school restructuring on high risk students.* Paper presented at the Annual Meeting of the American Educational Research Association, Chicago.

Aschbacher, P.R., & Herman J.L. (1990, September). *Issues in developing alternative assessments.* Presentation at the Annual Leadership Conference of the United Teachers of Los Angeles, Palm Springs, CA.

Baker, E.L., Aschbacher, P.R., Niemi, D., & Chang, S. (1991, February). *Alternative approaches to assessing students' deep understanding in social studies.* Los Angeles: UCLA Center for the Study of Evaluation.

Baker, E., Freeman, M., & Clayton, S. (1990). *Cognitive assessment of subject matter: Understanding the marriage of psychological theory and educational policy in achievement testing* (Tech. Rep. No. 317). Los Angeles: UCLA Center for the Study of Evaluation.

Linn, R., Graue, E., & Sanders, N. (1990). *Comparing state and district test results to national norms: Interpretations of scoring "above the national average"* (Tech. Rep. No. 308). Los Angeles: UCLA Center for the Study of Evaluation.

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform.* A report to the nation and the Secretary of Education. Washington, DC: U. S. Department of Education

Nickerson, R. (1989, December). New directions in educational assessment. *Educational Researcher, 18*(9), 3-7.

Resnick, L.B., & Klopfer, L.E. (1989). Chapter 1. Toward the thinking curriculum: An overview. In L.B. Resnick & L.E. Klopfer (Eds.), *Toward the thinking curriculum: Current cognitive research. 1989 ASCD Yearbook.* Alexandria, VA: Association for Supervision and Curriculum Development.

Table 1

Results of CRESST Survey: State Interest and Activity in Alternative Assessment 1990

States with Alternative Assessments in Place

CA	In place (math)—developing more math, science, social studies, writing
DE	In place (P.E., geography)—no other plans
HI	In place (math; alternative graduation exam in life skills)—exploring art, P.E.
ME	In place (reading, math)—interested in science
MA	In place (math, science, social studies, reading)
MI	In place (music, art, P.E.)—developing employability skills portfolios- have done science and career development in past—interested in social studies; tentative plans in science
NY	In place (science, math, social studies, second language, listening, speaking)
NC	In place (math)—interested in science, social studies, writing, speaking, second language, P.E.

States Currently Developing/Piloting Alternative Assessments

AK	Developing (portfolios in writing, interested in math)
AZ	Developing (reading, writing, math, social studies, science; subjects may be integrated; includes portfolios, pre-reading & pre-writing)
CT	Developing (math, science, writing, listening)
NJ	Developing (reading & math [open-ended])
VT	Developing (math & writing portfolios; encouraging schools to include writing across curriculum in portfolios; planning science, history, citizenship)

States Currently Exploring Possibilities (via committees or staff)

AL	Exploring (math)
CO	Exploring (math, science)
IL	Exploring (math)—exploring working with districts on science; interested in social studies
IN	Exploring (science, math, social studies)
MD	Exploring (RFP for creative, integrative approaches to whole language maybe w/social studies, science, math
MN	Exploring/Developing (science, social studies, three-day writing)
NM	Exploring (reading, writing portfolios)—voluntary participation
OR	Exploring (writing portfolios, math)—interested in P.E., fitness/health, science, social studies, art, music
KY	Plans (curricular goals not set as of spring 1990)
TX	Plans (interested in integrated writing & social studies)

* Based on phone interviews with state testing officers and staff; includes various alternatives to multiple-choice testing but does not include writing assessments unless they involve portfolios, writing across the curriculum, or other innovative approaches.

Table 1 (continued)

States Not Currently Involved in Alternative Assessment

AR	No plans (interested in writing portfolios)
FL	No plans (interested in science portfolios)
GA	No plans (interested in science)
IA	No statewide testing
ID	No plans (committee next year to consider possibilities & start developing)
KS	No plans (working with districts on math portfolios)
LA	No plans
MS	No plans
MO	No plans
MT	No plans
NE	No statewide testing (working w/teachers on writing, math, art portfolios)
NV	No plans
NH	No plans
ND	No plans
OH	No plans
OK	No plans
PA	No plans—(interested in science)
RI	No plans—(IN PLACE: unusual voluntary merit recognition program for 12th graders in art, vocational skills, & academic subjects—not part of state assessment)
SC	No plans (considering calculators in math)
SD	No plans
TN	No plans
UT	No plans (interested in writing in content areas)
VA	No plans
WA	No plans
WV	No plans
WI	No plans (interested in science, language arts)
WY	No plans for statewide (districts will do in language arts, math, etc. for new accreditation regulations)

Table 2
Survey of State Writing Assessment Programs
Fall 1990

States	AL	AK	AZ	AR	CA	C O	CT	D E	FL	GA
a. Test writing?	*Dev*	yes	dev	no (yes in 5yrs.)	yes	yes	yes	*yes*	*no*	yes
b. Type: multi-choice written	*w*	w			mc w	w	mc w	*w*		w
c. Grades	*2,5*	10			3,6,8,12 (chg to 4,8,11)	4,7,10	4,6,8 9-12	*10*		6,8,10
d. Sample all smpl matrix smpl	*S*	voluntary			all & ms	ss	ev	*all*		all
e. No. of writing samples		1			1		1			1
f. revisions assessed					no	no	no			no; pre-w not scored
g. time to write		20 days			45 min.	40 min.	40 min			2 hours
h. scoring: analytic, holistic primary trait		an			ho	pt	an, ho			an, ho
i. who developed	*cont.*	state			state	state	state	*state*		state
j.who scores	*cont*	Teachers			Ts	Meas Inc	Ts	*cont.*		Ts & others
k. train teachers		yes			yes	yes	yes			no
l. test wrtg in subjs (future)					hist, soc st, math, (sci)	(his, sci, math)	(his, sci, math)			

Table 2 (Cont'd.)

States	HI	ID	IL	IN	IA	KS	KY	LA	ME	MD
a. Test writing?	*yes*	yes	yes	yes	*opt'l (no st. ass'mt)*	no (yes in 5 yrs.)	*No (yes in 5 yrs.)*	yes	*yes*	yes
b. Type: multi-choice written	*w*	w	mc w	w	*w*		*w*	w	*w*	w
c. Grades	*9-12*	8,11	3,6,8,11	3,6,8,9,11	*3-8*		*3,8,12*	5,7,10 7,10	*4,8,11*	9
d. Sample all smpl matrix smpl	*vol*	all	all	all	*vol.*		*s*	gr 7,10 = all gr 5=5	*all*	all
e. No. of writing samples		1	1	1				1		2
f. revisions assessed		no	no	no				yes		yes
g. time to write		90 min	45 min	32-45 min				60 min		untimed
h. scoring: analytic, holistic primary trait		ho	an, ho	an, ho				domain scoring model		ho
i. who developed	*state*	scoring team	state	state, CTB	*state*			consul-tant	*cont.*	state
j.who scores	*Ts*	Ts	Meas Inc	CTB	*Ts*			Meas Inc	*Ts & others*	Meas Inc
k. train teachers		yes	yes	no				yes		yes
l. test wrtg in subjs (future)		his, sci, (math)		none						ma (his, sci)

Table 2 (Cont'd.)

States	MA	MI	MN	MS	MO	MT	NE	NV	NH	NJ
a. Test writing?	yes	no (yes in 5 yrs.)	yes	*yes*	yes	no (yes in 5 yrs.)	*no (no st. ass'mt)*	yes	*no*	dev
b. Type: multi-choice written	w	w, mc	mc; w	*w*	mc; w			mc; w		w
c. Grades	3,6,9	4,5,7,8, 10,11	6,9,11	*8*	8			mc=3,6 w=9,11, 12		8
d. Sample all smpl matrix smpl	all	s	ms	*all*	all			all		
e. No. of writing samples	1	1	1		1			2		
f. revisions assessed	yes	yes	no		yes			no		
g. time to write	untimed	(30 min, 3 days)	3 hr/ 3 days		1 hr- 1 wk			30		
h. scoring: analytic, holistic primary trait	ho	ho	ho		ho			ho		
i. who developed	state	NAEP/ state	state	state	state			state		CTB
j.who scores	Ts	Ts	Ts	Ts	Ts			Ts		CTB
k. train teachers	yes	yes	yes		yes			yes		
l. test wrtg in subjs (future)		(his, sci, math, read'g)								

Table 2 (Cont'd.)

States	NM	NY	NC	ND	OH	OK	OR	PA	RI	SC
a. Test writing?	*dev (opt)*	yes	*yes*	no	yes	yes & dev	yes & dev	dev	yes	yes
b. Type: multi-choice written	*w*	w			w	mc,w	w	w	w	w
c. Grades	*10-12*	5,8,9,11	*6,8*		9	7,10	3,5,8,11	6,9	3,6	6,8,10-12
d. Sample all smpl matrix smpl		all	*all*		all	all	all	ms	all	all
e. No. of writing samples		2-3			2	1	1	1	1	1
f. revisions assessed		no			yes	no	no	no	no	no
g. time to write		1 hr			2.5 hrs	20 min	3 cl per.	100 min	90 min	–
h. scoring: analytic, holistic primary trait		ho			an	an, ho	an	ho	ho	an, ho
i. who developed	*state*	state	*Ts*		state + m.i.	Psy Corp	state (CUNY)	Adv Sys	state	state
j.who scores	*Ts*	Ts & others	*Ts*		Meas Inc	Psy Corp	Ts	s Ts	Ts	DRC
k. train teachers		yes			no	yes	yes	yes	yes	yes
l. test wrtg in subjs (future)		his					(math)			(read'g)

Table 2 (Cont'd.)

States	SD	TN	TX	UT	VT	VA	WA	WV	WI	WY
a. Test writing?	no	dev (opt)	yes	dev (opt)	yes	yes (opt)	yes (opt)	yes	opt (past yes)	no
b. Type: multi-choice written		mc, w	mc, w	w	w	w		w	mc, w	
c. Grades		4,8,11 s	3,5,7,9.11	5,8,11 ev	4,8	K-12		8,10	8,10	
d. Sample all smpl matrix smpl		s	all	all	s			all	all	
e. No. of writing samples		1	1	opt'l	portfolio			1	1-3	
f. revisions assessed		no	no	no				no	no	
g. time to write		25 min	untimed					60 min	30-45 min	
h. scoring: analytic, holistic primary trait		ho	an, ho	an				ho	ho, pt	
i. who developed		state + ETS	state Meas. Inc.	NWREL	Ts	state		state	state	
j.who scores		CTB	Meas Inc	Ts	Ts	Ts		Ts	Ts	
k. train teachers		yes	yes	yes				yes	yes	
l. test wrtg in subjs (future)			(his, sci)	(his, sci)						

Appendix A

Survey of State Writing Assessment Programs
Fall, 1990

In the fall of 1990, CRESST surveyed the testing directors in each of the 50 states about their writing assessment programs and future plans. The writing survey was contained within a longer survey conducted annually by the Association of State Assessment Programs (ASAP) of its members (a copy of the survey questions follows on the next page). Thirty-three states responded, and the information they provided is displayed in Table 2 (in regular print). This table also includes some data (in italics) on states that did not respond to the ASAP survey. These data are drawn from a recent survey conducted by the Chief State School Officers[1] on performance assessments in the states.

A perusal of Table 2 indicates the following:

- 28 states currently have a statewide direct writing assessment;
- 3 have direct writing assessments as a district option;
- 7 are currently developing direct writing assessments;
- 6 expect to begin development within five years; and
- 6 states do not have any plans to assess direct writing in the near future.

A few states (7) collect more than one writing sample per student per year, and quite a few (12) allow students more than an hour to write. None of the states currently assesses writing in cooperative group settings, although Manitoba, Canada does. Very few states (4) now assess students' writing in subject matter areas such as history, math or science, but several (7) indicated interest in doing so in the future. The great majority of states with direct writing assessments involve teachers in scoring students' work (25), and most of them provide the associated training for teachers.

Several states noted that they would like technical assistance related to direct writing in the following areas:

- good prompt ideas and characteristics of good prompts,
- equating of prompts across years,
- use of writing to assess subject matter areas,
- use of portfolios,
- comparability of scoring systems,
- procedures for setting cut scores,
- relationship between writing time and performance, and
- use of word processors to assess writing.

---

[1] Council of Chief State School Officers. (1990, December). *Performance Assessments in the States.* Presentation to the Assessment Committee of the Secretary's Commission on Achieving Necessary Skills. Washington, DC: Pelavin Associates, Inc.

Writing Assessment Activities

A. Do you currently assess writing or are considering doing so?

_____ yes, currently test writing
_____ have writing test in development
_____ no, but plan to within 5 years
_____ no, and do not plan to in near future
_____ other (describe) _____

B. Which of the following do you use? (Check all that apply)

_____ multiple-choice items
_____ Standardized test (which?) _____

_____ Other (who developed?) _____

_____ writing sample(s) in response to standard prompt(s)
_____ Standardized test (which?) _____

_____ Other source (who developed?) _____

_____ samples of existing work (portfolio)
Who decides which work is included? _____

C. Grades tested: 1 2 3 4 5 6 7 8 9 10 11 12

D. _____ every pupil testing _____ simple sample _____ matrix sample

E. Number of writing samples per student per year? _____

F. Are revisions assessed? _____ yes _____ no

G. Length of time student has for writing sample _____

H. How do you score? ____ analytic ____ holistic ____ other

I. Who developed the scoring rubric? _____

J. Who scores writing samples?
_____ teachers in state
_____ commercial company (which?) _____
_____ other (who?) _____

K. Do you provide related teacher training? _____ yes ____ no

L. Check content areas in which you assess writing or consider doing:

Currently test writing in:	Considering testing writing in:
_____ history, social studies	_____ history, social studies
_____ science	_____ science
_____ math	_____ math

M. Do you currently assess writing in cooperative group settings?
_____ yes ____ no _____ considering in future

N. What technical assistance or research information related to writing would be helpful to you? _____