C-R-E-S-S-T

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

Technical Report

You can view this document on your screen or print a copy.

- UCLA Center for the Study of Evaluation

in collaboration with:

- University of Colorado
- NORC, University of Chicago
- LRDC, University of Pittsburgh
- The RAND Corporation

Parent Opinions About Standardized Tests, Teacher's Information and Performance Assessments

A Case Study of the Effects of Alternative Assessment in Instruction, Student Learning and Accountability Practices

CSE Technical Report 367

Lorrie A. Shepard and Carribeth L. Bliem
CRESST/University of Colorado at Boulder

October 1993

National Center for Research on Evaluation,
Standards, and Student Testing (CRESST)
Graduate School of Education
University of California, Los Angeles
Los Angeles, CA 90024-1522
(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.

PREFACE

The current intense interest in alternative forms of assessment is based on a number of assumptions that are as yet untested. In particular, the claim that authentic assessments will improve instruction and student learning is supported only by negative evidence from research on the effects of traditional multiple-choice tests. Because it has been shown that student learning is reduced by teaching to tests of low level skills, it is theorized that teaching to more curricularly defensible tests will improve student learning (Frederiksen & Collins, 1989; Resnick & Resnick, 1992). In our current research for the National Center for Research on Evaluation, Standards, and Student Testing (CRESST) we are examining the actual effects of introducing new forms of assessment at the classroom level.

Derived from theoretical arguments about the anticipated effects of authentic assessments and from the framework of past empirical studies that examined the effects of standardized tests (Shepard, 1991), our study examines a number of interrelated research questions:

1. What logistical constraints must be respected in developing alternative assessments for classroom purposes? What are the features of assessments that can feasibly be integrated with instruction?

2. What changes occur in teachers' knowledge and beliefs about assessment as a result of the project? What changes occur in classroom assessment practices? Are these changes different in writing, reading, and mathematics, or by type of school?

3. What changes occur in teachers' knowledge and beliefs about instruction as a result of the project? What changes occur in instructional practices? Are these changes different in writing, reading, and mathematics, or by type of school?

4. What is the effect of new assessments on student learning? What picture of student learning is suggested by improvements as measured by the new assessments? Are gains in student achievement corroborated by external measures?

5. What is the impact of new assessments on parents' understandings of the curriculum and their children's progress? Are new forms of assessment credible to parents and other "accountability audiences" such as school boards and accountability committees?

This is one of four reports that document our progress in understanding these questions, based on case studies in three elementary schools.

PARENT OPINIONS ABOUT STANDARDIZED TESTS, TEACHER'S INFORMATION AND PERFORMANCE ASSESSMENTS[1]

Lorrie A. Shepard and Carribeth L. Bliem
CRESST/University of Colorado at Boulder

New forms of assessment are seen by many as the centerpiece of current educational reform efforts. A growing body of research has documented the negative effects of standardized testing on teaching and learning (Madaus, West, Harmon, Lomax, & Viator, 1992; Shepard, 1991; Smith, 1991). Traditional forms of testing—filling in the blanks and picking right answers—have contributed to an implicit national curriculum focused on low-level skills (Resnick & Resnick, 1992). At the same time, efforts to gain external control over classrooms via standardized measures and teacher-proof curricula have led to the deskilling and deprofessionalization of teachers (Darling-Hammond, 1988; McNeil, 1988). In contrast, reformers expect performance assessments to have a nondistorting and positive effect on instruction and learning by aiming at more direct, complete, and integrated measures of desired student outcomes.[2]

Although there is widespread enthusiasm among educators for portfolios and other forms of performance assessment, implementation is problematic for both technical and political reasons. In this paper we address one formidable obstacle—the objection of parents.

Policymakers perceive a demand for standardized testing from both parents and the general public. Standardized test score results are deemed essential to monitor individual student progress and to hold schools accountable. This demand is evidenced by annual Gallup Poll results demonstrating the high percentage of respondents favoring the use of standardized national tests by public schools in their community to measure the academic achievement of

---

[1] This paper was presented at the annual meeting of the American Educational Research Association, Atlanta, GA, April, 1993.

[2] In this paper we use the term performance assessment as defined by Resnick and Resnick (1992). "A performance assessment is one that uses direct judgments and evaluations of performances rather than indirect indicators of competence" (p. 61). We consider performance assessments to be synonymous with *direct* assessment as described by Frederiksen and Collins (1989) and *authentic* assessment described by Wiggins (1989).

students (Elam, Rose, & Gallup, 1992). As a result, legislators, school board members, and professional educators fear that parents will be "up in arms" if performance assessments are adopted to supplant familiar standardized tests. In the early stages of the current research project, for example, sympathetic members of the district accountability committee asked that researchers discontinue the use of the phrase "alternative assessment" because it connoted abandonment of standards and rigor.

The purpose of this study was to examine in greater depth parent opinions about standardized tests and performance assessments. The research was guided by the following questions: (a) How do parents in this sample respond to marker Gallup Poll questions about the quality of local schools, the quality of *information* provided to parents, the desirability of standardized national tests, and the potential uses for standardized test results? (b) How do parents evaluate the usefulness of standardized tests compared to less formal types of information such as report cards, talking to the teacher, or seeing samples of their child's work? Do parents differentiate between information they find useful in learning about their child's progress in school versus the information they use to make judgments about the school as a whole? (c) After being shown sample questions and tasks, what preferences do parents express regarding district use of standardized tests and performance assessments in reading and mathematics? Is it true that parents disdain the use of performance assessments as less rigorous or objective? What features of the two types of measures affect parent preferences? (d) How do parents respond when asked to make their judgments about standardized test questions and performance assessments in the context of classroom instruction?

Methods

Larger Research Project

The present study of parent opinions was conducted in the context of a larger research and development project. In the larger project, researchers worked with third-grade teachers for the entire 1992-93 school year to develop classroom performance assessments in reading and mathematics. The central aims of the research were to examine the effects of implementing performance assessments on instruction and student learning. Papers by Borko, Flory, and Cumbo (1993), Flexer and Gerstner (1993), and Hiebert and Davinroy (1993)

consider changes in teachers' thinking about assessment and instruction using beginning-of-year and mid-year interview data and transcript analysis from weekly workshop sessions. Effects on student achievement will be assessed at the end of the year by comparing third graders in the three participating schools to those in three control schools on the Maryland Assessments for reading and mathematics. (The accuracy of this quasi-experimental comparison is strengthened by matching of schools on socio-economic variables, baseline comparisons of the two sets of schools on the Maryland Assessments in May 1992 before the start of the study, and covariate adjustments based on premeasures administered to third graders in September 1992.) A collateral study also addresses changes in students' perceptions of teacher evaluations during the course of the project.

Parent Questionnaire and Interview Samples

The study was conducted in a working-class and lower-middle-class school district on the outskirts of Denver, Colorado. The district was selected on the basis of three criteria: (a) an ethnically diverse student population, (b) a history of standardized accountability testing, and (c) a willingness on the part of district personnel to seek a two-year waiver from standardized testing in the three schools selected to participate.

The three participating schools were selected because the principals and third-grade teams of teachers volunteered. As part of the formal process required by the Colorado Department of Education to apply for a waiver from standardized testing, approval was sought from each school's accountability committee. Thus, although data on teachers' beliefs and practices suggest that teachers were fairly traditional in their approach to classroom assessment, the volunteer nature of the school samples makes it probable that the school communities in this study were slightly more favorably disposed toward performance assessments than are schools in general.

Three control schools were also identified to be used for comparative purposes when analyzing student achievement, teacher beliefs, and parent opinion. The control and participating schools were matched on free and reduced lunch percents, percent minority, and other knowledge of neighborhood similarities. The following display illustrates socio-economic differences among the three participating schools as well as their matches to control schools.

	Participating schools			Control schools		
	1	2	3	1	2	3
Free and reduced lunch	61%	9%	6%	55%	13%	3%
Percent minority	37%	16%	14%	45%	19%	10%

Questionnaire parent samples were selected in both the participating and control schools. Interview parent samples were identified only in the participating schools. The sampling design was developed such that mutually exclusive but randomly equivalent stratified samples of parents could be used to measure change in parent opinions over time. Data reported in this paper, however, represent only the premeasures collected in the fall at the end of the first academic quarter. Questionnaires were sent to a random one-third of parents in each participating and control classroom. Overall, response rates were 70% for the participating schools and 49% for the control schools, resulting in samples of 69 and 36 respectively.

In the participating schools, parents of three children per classroom were sampled to be interviewed in conjunction with first-quarter parent conferences. Children were stratified by achievement level (high, middle, and low). Teachers were provided with two alternate names within each stratum to use in cases where parents declined to participate or where scheduling precluded participation. We have no accurate way to estimate what biases were introduced into the sample by self-selection or schedule conflicts. However, it was a relatively rare occurrence for teachers to pick the second name on the list because a conference was going to be problematic or because a parent was thought to be unresponsive to school requests. Ironically, a greater percentage of parents were willing to commit to the time demands of the interview than returned questionnaires. After the fall interview sample was agreed upon (n = 13 classrooms X 3 = 39), attrition occurred in the sample because of absence from the conference (n = 5) and technical failures in recording (n = 1), resulting in a final interview sample of 33 parents or parent dyads. Absence from the scheduled parent conferences was related to student achievement level; 3 of the absent parents had children in the low ability group and 2 were from the middle group. Therefore, the final sample was not as balanced as originally intended.

Questionnaire and Interview Instruments

The four-page mailed questionnaire had three parts. The first part asked parents to rate the usefulness of report cards, talking to the teacher, standardized tests, and seeing graded samples of their child's work. The second part provided displays showing typical questions on third-grade standardized achievement tests and typical performance assessment tasks in both reading and math. Parents were then asked to rate their degree of approval for each of the four types of measures. Then they were asked to show their preference for one or the other type of measure when used in instruction. The third set of questions was taken from the 1992 Gallup Poll. Parents responded to the questionnaires anonymously.

Interviews, conducted immediately following parent conferences, were guided by a structured but open protocol. The interview was divided into three parts, two of which closely paralleled the parent questionnaires. Parents were first asked broad questions aimed at eliciting their conceptions of reading and mathematics knowledge and what they used as indicators to judge their child's progress in each area. Two forms of the interview were printed so that half of the interviews began with reading questions and half began with mathematics. The second set of questions asked parents to rate the usefulness of report cards, talking to the teacher, standardized tests, and seeing graded samples of work. Parents were asked to think aloud and elaborate on their answers while filling out the rating form. For the last set of questions, parents were shown more extended examples of questions from standardized tests and performance assessments in reading and mathematics. They were asked about their approval or disapproval of each and about the advantages or disadvantages they saw with each type of measure. A final question asked parents which of the two kinds of tests they would like to see used in classroom instruction.

Interviews were audio recorded and transcribed. Extensive field notes were also taken. In one case where a parent declined to be recorded, answers were typed directly into a computer.

Findings

Gallup Poll Results

Nationally 71% of the general public as well as 71% of public school parents favor requiring public schools in their community to use standardized national tests (Elam et al., 1992). If the purpose of the present study is to go beyond this superficial endorsement of standardized tests and develop a deeper understanding of parent opinions, then it is important to know how similar parents in the study sample are to the national population of parents.

Table 1 provides a summary of questionnaire results for third-grade parents in the study compared to data from the 1992 Gallup Poll (Elam et al., 1992). A preliminary analysis showed no significant differences between questionnaire responses for control and participating schools, therefore they are combined here as a single sample.

On the most classic Gallup question (#21) concerning national tests, there is a disconcertingly large disparity between third-grade parents in this study and the national sample of public school parents. Only 46% of local third-grade parents favor national tests compared to 71% nationally. Opposition rates are more similar because a very large proportion of third-grade parents (25%) responded in the Don't-Know category. On other questions, the gap between the two groups is much narrower suggesting that the third-grade sample can be thought of as only slightly more negatively disposed toward standardized testing than the nation as a whole. On questions 22-27 addressing various uses of standardized national tests, the percentage of third-grade parents endorsing each use was consistently 12% to 14% less than for the national group, with the exception of using test results to improve teaching where the difference was only 7%.

On the more general Gallup questions dealing with quality of schools and quality of information provided by schools about each child's academic progress, the third-grade parents rated their schools higher (66% gave As and Bs compared to 40% for public school parents nationally); they rated the nation's schools lower (9% As and Bs vs. 19% nationally); and they rated the quality of information they receive just the same as did a national sample of elementary parents (69% vs. 71% As and Bs).

Table 1

A Comparison of National Results and Third-Grade Parent Responses (n=105) on 1992 Gallup Poll Questions

18. Students are often given the grades A, B, C, D, and Fail to denote the quality of their work. Suppose the public schools themselves, in this community, were graded in the same way. What grade would you give the public schools here—A, B, C, D, or FAIL?

	A	B	C	D	FAIL	Don't Know
1992 National %	9	31	33	12	5	10
3rd-Grade Parent Sample %	13	53	25	2	2	5

19. How about the public schools in the nation as a whole? What grade would you give the public schools nationally—A, B, C, D, or FAIL?

	A	B	C	D	FAIL	Don't Know
1992 National %	2	16	48	18	4	12
Public School Parents National %	2	17	48	18	4	11
3rd-Grade Parent Sample %	1	8	50	20	7	15

20. What grade would you give the information you receive from your third-grade child's teacher(s) regarding his or her academic progress—A, B, C, D, or FAIL?

	A	B	C	D	FAIL	Don't Know
1992 National %	31	30	23	7	6	3
Public School Parents National %	42	29	20	6	2	1
3rd-Grade Parent Sample %	32	37	20	6	1	4

21. Would you favor or oppose requiring the public schools in this community to use standardized national tests to measure the academic achievement of students?

	1992 National %	Public School Parents National %	3rd-Grade Parent Sample %
Favor	71	71	46
Oppose	20	23	30
Don't Know	9	6	25

In addition to measuring the academic achievement of students, do you think these standardized national tests should be used or should not be used for the following purposes?

	1992 National %			3rd-Grade Parent Sample %		
	Should	Should Not	Don't Know	Should	Should Not	Don't Know
22. rank public schools	65	26	9	53	37	10
23. grade promotion	60	32	8	46	45	9
24. teacher pay	38	52	10	26	62	12
25. school funding	36	54	10	22	68	10
26. improve instruction	79	14	7	72	17	11
27. student weaknesses	85	9	6	73	19	8

Given the above similarities and dissimilarities between national samples of parents and the third-grade sample here, we suggest that the findings in this study—about broad parent preferences and especially about reasons parents give for their preferences—are likely to be generalizable beyond the confines of this particular district. However, it is unlikely that strict quantifications, such as the percent preferring performance assessments over standardized tests, can be generalized either to nonrespondents in the same district or to the nation as a whole.

Parent Opinions About Standardized Tests and Informal Sources of Information: Questionnaire Results

On the questionnaire, parents were asked to evaluate the usefulness of four types of information for two different purposes, learning about their child's academic progress and judging the quality of the school. Data in Table 2 show the percentage of responses for the first purpose. Overwhelmingly, parents believe that they learn the most about their child's progress by listening to the teacher; 77% rated this source of information as very useful. A total of 94% rated hearing from the teacher in the top two categories. All three informal sources of information about student progress—report cards, talking to the teacher, and work samples—received strikingly higher usefulness ratings than did standardized tests. However, it should be noted that a biasing factor in this study is that these third-grade parents were surveyed in the fall before the first opportunity for their third-grade child to be tested. Therefore, parents were likely to be responding on the basis of experience with older children or a general opinion they held regarding standardized tests. Parents who had little knowledge of standardized tests and whose oldest child was in third grade may not have had sufficient information to judge the usefulness of such tests.

In Table 3 parent ratings are presented for the same four types of information but for the purpose of judging the quality of the school. As expected, more parents considered standardized tests to be useful for this purpose than for learning about the progress of individual children (45% vs. 36%). In addition, 19 respondents (18%) wrote in comments suggesting that it is the comparative information provided by such tests that makes them useful for evaluating the school's performance.

Table 2

Parent Questionnaire Ratings of the Usefulness of Each Type of Information for Learning About Their Child's Progress in School (n=105)

Place a check in the appropriate column to indicate the usefulness of each type of information.	Not at all useful				Very useful	Blank/ Missing
	1	2	3	4	5	
1. Report cards	2%	2%	20%	33%	43%	
2. My child's teacher talking about his/her progress	0	2%	4%	17%	77%	
3. Standardized tests	6%	15%	41%	22%	14%	2%
4. Seeing graded samples of my child's work	0	0	10%	30%	60%	

5. Pick one type of information you think is most helpful in learning about <u>your child</u> and tell what you learn from it and why it is helpful.

Report cards (n=5)

"I think the most important item is report cards because it shows how well they are doing in each subject and then we can determine where help is needed."

Teacher talking (n=44)

"Talking with my child's teacher is most helpful because I learn firsthand what progress is being made in class, where the shortcomings are and how I can best help at home."

Standardized tests (n=0)

Graded samples (n=38)

"This way I can see the actual work, the teacher's response, and evaluate what I understand the child's level of learning to be at."

"When we are able to see our child's work we can see the progress that he makes and can help him in the areas that he needs help."

Other (n=9)

However, much of the data in Table 3 are contrary to the expectation that parents would prefer external and formal measures for school accountability purposes. Indeed the greatest number of parents, 84% and 83% respectively, valued "my child's teacher talking about his/her progress" and "seeing graded

Table 3

Parent Questionnaire Ratings of the Usefulness of Each Type of Information to Evaluate the Quality of Education Provided at Their Child's School (n=105)

Place a check in the appropriate column to indicate the usefulness of each type of information.	Not at all useful				Very useful	Blank/ Missing
	1	2	3	4	5	
6. Report cards	18%	11%	35%	17%	16%	3%
7. My child's teacher talking about his/her progress	4%	2%	18%	28%	46%	3%
8. Standardized tests	8%	14%	30%	20%	25%	4%
9. Seeing graded samples of my child's work	1%	4%	20%	27%	46%	3%

10. Pick one type of information you think is most helpful in learning about <u>the quality of education provided at your child's school</u> and tell what you learn from it and why it is helpful.

Report cards (n=0)

Teacher talking (n=26)

"Talking with the teacher allows you into the school environment to measure emphasis, attitudes, some insight into treatment of students (i.e., respect, encouragement, etc.)."

"Again I would say talking with the teacher in person. That way I get an understanding of the school policies and programs and what curriculum the school uses."

Standardized tests (n=19)

"Standardized tests show me where [my school] ranks in comparison with district, state and national averages —"

Graded samples (n=18)

"Again, samples of my child's work show me what it is they are doing in the classroom, and will show me where the school's emphasis is at."

Other (n=25)

"Newsletters from school sent home with the student. These discuss what the goals are that have been established, what type of work the students are doing, what teachers expect, etc."

"Being able to be in the classroom, getting to know the teacher, seeing hands-on, spending time and helping at school."

samples of my child's work" to help them evaluate how good a job the school was doing. Furthermore, in follow-up comments parents gave justifications that showed they were mindful of the change in purpose from the first set of questions to the second. Parents say that these informal sources help them learn about the quality of education by giving them first-hand information about the school curriculum, teacher knowledge and expectations, school policies, and classroom climate.

Given earlier differences found between the study sample and the national population on Gallup questions, we were concerned that the above, somewhat surprising findings—downplaying the usefulness of standardized tests—might be true only for parents in this sample and not generalize to a broader population of parents. Therefore the present sample was subdivided on the basis of Gallup Poll question 21 into those parents who "Favor" use of standardized national tests and those who "Oppose" (parents answering Don't Know were omitted). Predictably these two groups differed in their ratings of the usefulness of standardized tests on Question 3 in Table 2. However, they did not differ at all in their much higher ratings of informal sources of information. For example, on Question 2, teacher talking, the mean ratings were 4.65 and 4.74 for the Favor and Oppose groups respectively. For Question 4, seeing graded samples of work, the means were 4.58 and 4.45. (Neither difference is statistically significant and the slight numeric differences are offsetting.) On the second set of questions dealing with accountability purposes, the two groups again differed significantly in their usefulness ratings for standardized tests ($\overline{X}$ Favor = 3.94; $\overline{X}$ Oppose = 2.80). They did not differ in their ratings on the other three sources of information; both groups gave report cards low marks for this purpose but rated teacher talk ($\overline{X}$ Favor = 4.02; $\overline{X}$ Oppose = 4.10) and work samples ($\overline{X}$ Favor = 4.26; $\overline{X}$ Oppose = 4.00) highly. This means that the group favorable to standardized testing, which may represent a greater proportion of the national population, only considered standardized tests to be as useful as talking to the teacher for accountability purposes and not as useful as seeing graded work samples. For the purpose of learning about an individual child's progress, both groups agree strongly that all three sources of information are more useful than standardized tests.

Parent Opinions About Standardized Tests and Informal Sources of Information: Interview Results

In the interview data, parents reported even more glowing opinions about the usefulness of informal types of information for learning about their child's academic progress than they had on the questionnaire. In Table 4 and Table 5, quantitative results are presented from the interview questions corresponding to the questionnaire data shown in Table 2. The question about standardized tests was pulled out and treated in a separate table to allow for separate analysis of the qualitative data. Although the parent responses about the usefulness of standardized tests are virtually the same (36% of the questionnaire sample said they were useful vs. 33% in the interview sample), the interview respondents were even more positive about the usefulness of report cards, hearing the child's teacher talk about progress, and graded samples of work than the questionnaire respondents had been. For example, 77% had rated teacher talk as "Very useful" on the questionnaire, whereas 91% of the parents interviewed gave the "Very useful" rating. Because parents in the interview setting were asked to fill out a rating sheet identical to the question set used on the questionnaire, it is unlikely that the demand characteristics of the interview (having to explain their answers, etc.) account for the more extreme answers in the interview. The more likely explanation is that parents were interviewed immediately following parent conferences. It is also possible that the interview

Table 4

Parent Interview Ratings Regarding the Usefulness of Informal Types of Information for Learning About Their Child's Progress in School (n=33)

How useful is each type of information listed below in helping you learn about <u>your child's progress in school</u>?

	Not very useful				Very useful	Blank/
	1	2	3	4	5	Missing
a. Report cards	0	3% (n=1)	18% (n=6)	15% (n=5)	61% (n=20)	3% (n=1)
b. Talking with the teacher	0	0	0	6% (n=2)	91% (n=30)	3% (n=1)
d. Seeing graded samples of my child's work	0	0	0	15% (n=5)	82% (n=27)	3% (n=1)

Table 5

Parent Interview Responses Regarding the Usefulness of Standardized Tests for Learning About Their Child's Progress in School (n=33)

How useful is each type of information listed below in helping you learn about <u>your child's progress in school</u>?

	Not at all useful				Very useful	Blank/
	1	2	3	4	5	Missing
c. Standardized tests	12% (n=4)	21% (n=7)	27% (n=9)	21% (n=7)	12% (n=4)	6% (n=2)

Normative information n=4 ($\bar{X}$=4.7)

"It compares him to everybody else and lets you know how he's doing compared to them. Because if he's just getting As on everything and everybody else is getting As for not doing anything at all, then it doesn't say much for him."

"How they do compared to other children is very important so that you have an idea that they are progressing like they should."

Helpful academically n=4 ($\bar{X}$=4.2)

"I think they are very useful. Not to me but to the teachers. I think they probably should have these so the teachers can inform the parents as to where they're at. Informs me as to what she needs help on."

Other positive attributes n=3 ($\bar{X}$=4.0)

Children as individuals n=6 ($\bar{X}$=3.0)

"Some children just do not respond well to certain standardized tests. I think that children are different types of learners and standardized testing sometimes does not always catch that, the full breadth of a child's learning abilities."

"A standardized test really doesn't center around what the problems are with the individual child, but are used as a measurement of how much this child knows."

Invalid results n=6 ($\bar{X}$=2.7)

"I think it's got a lot to do with how well a child takes tests, what their particular morning was like, or whether they had a good night's sleep. I think there's so many variables that can influence their score on the standardized testing."

Parents' own experiences n=2 ($\bar{X}$=1.5)

"I don't like them. And I probably have to say because when I was a kid they made me nervous. I do better without the test. And when they go to give you that big test, it's overwhelming and you know you're being graded on whether you're stupid or smart."

Don't know n=2 ($\bar{X}$=2.5)

Child hasn't taken test or parents have never seen results n=3 ($\bar{X}$=2.0)

Other negative attributes n=2 ($\bar{X}$=1.0)

Unfamiliar with standardized tests n=1 (parents left question blank)

and questionnaire data differ because they are based on different samples; the interview data were collected only in participating schools as compared to the questionnaire data which were collected from both control and participating school parents. However, there were no significant differences in the questionnaire results between the two samples on these questions.

Transcript analyses of parent comments explaining their ratings are summarized in Table 5 and Table 6. In Table 5, parent comments about why standardized tests are useful or not useful are related predictably to their ratings (means are shown in parentheses). Parents who rated standardized tests highly commented on the usefulness of normative comparisons and the ability of tests to identify academic strengths and weaknesses. Parents who gave lower ratings to standardized tests tended to comment on the vagaries of test performance on any given day or their own negative experiences taking standardized tests.

Parents had a great deal to say about why they valued informal sources of information, especially talking to the teacher, to learn about their child's progress. Note that all three informal types of information, even report cards which got the lowest rating, were rated as more useful than standardized tests. Parents want to know in some detail how their child is doing. They want to know strengths and weaknesses and what they can work on at home. To interpret the meaning of where their child is, they seek comparative information, that is, how their child is doing in relation to grade level expectations or compared to the rest of the class.

Perhaps the most interesting finding, however, is that parents, by and large, trust their child's teacher to provide them with this type of information. "She's the trained professional, she knows what to look for . . . " "She deals with my child every day. She knows my daughter and can tell me things about her learning." We considered the very high approval ratings given to teacher information to be striking because usually the public endorsement for standardized national tests is taken to mean that parents require external verification of progress and cannot rely on teacher judgment.

Table 6

Parent Interview Responses Regarding the Usefulness of Informal Types of Information for Learning About Their Child's Progress in School (n=33)

Nearly three-quarters of respondents (73%, n=24) appreciate how these types of information provide feedback about how well their child is doing academically. They cite understanding their child's strengths and weaknesses as well as seeing evidence of his or her progress over time.

"If I didn't have the report card, there would be no way I'd be able to judge that he was actually being able to retain what he's learning. She [the teacher] shows me what I really need to work on with him and with him being there, he sees that he needs to work on this."

"The report cards are very useful. You can [see by the] quarterly report the improvement she's made or if she's not making improvement, you can tell."

"As long as she [the teacher] can see in math that he's improving with each of his tests, that's all she cares about. It's just important that he is doing the best he can do, and that he's improving."

"I like [graded samples]. I think it's useful because then I know from [the teacher] showing from September to now, I know that she's improving."

Over one-third of the respondents (36%, n=12) comment on the normative aspects of this information. They depend on the teacher to provide the comparative information.

"I think report cards are very useful because their pluses and minuses [system of grading] tell you where they're at in reference to the rest of the class basically."

"[Graded samples] give you more of a clear idea of where the child stands within the class itself."

"Well, [these types of information] let me know that they are doing their level of work. I know that they're at the level that they should be doing. That assures me that they are doing the level that they're supposed to be."

"Well, I think on the report cards, it's very useful because in her reading you know what level she is reading."

With regard to talking to the teacher, respondents value the opportunity for communication with the teacher (18%, n=6).

"If there's questions, just keeping communication with the teacher is important. If you have questions, sometimes things come home that you don't understand and you need to know the answers."

"Talking with the teacher, I think, is most important. We can then ask questions as to the report card . . ."

Talking with the teacher helps 24% of the parents (n=8) because the teacher spends time with their child.

"They [teachers] see what work they're doing more than we do because we both work. They're with our kids 9 hours a day."

Table 6 (continued)

"She's the trained professional, she knows what to look for if something should come up that we should be aware of . . . To be able to talk with somebody who can see their development and be there all times is very important."
"She deals with my child every day. She knows my daughter and can tell me things about my daughter's learning that she notices. She sees her in a context where she's maybe one-on-one at times, she's in a small group, she's in a large group, and sees my daughter in a wide variety of experiences, where I see her only in a family setting."
"They can see everything that they're doing, and they can see if they're messing around, not getting their assignments, or if they're struggling to understand . . ."

Note. Data from the three questions in Table 4 were treated as an aggregate. The themes emerged from the parent comments. Categories are not mutually exclusive.

When the purpose shifts from "learning about my child's progress" to evaluating the school, the interview results parallel the questionnaire findings but again with more extremely positive ratings for informal sources of information from the interviewed parents. Quantitative ratings are shown in Table 7 and Table 8. Just as with the questionnaire data, the usefulness of standardized tests increased when the purpose was to judge the school rather than an individual child's progress. On the questionnaire, the increase in positive ratings was from 36% to 45%; for the interview sample, the increase was from 33% to 45%. Again on the question about standardized tests, there was a close correspondence between the responses from the questionnaire and interview samples. In explaining their ratings, approximately one-third of the parents interviewed said something about the value of standardized tests for making comparisons. "That is one of the reasons I like the standardized tests, because to me if you have a national standard test for third graders, it shows you where your kid is against national standards which doesn't necessarily say anything about your kid but it might point out there is a problem here . . . "

Table 7

Parent Interview Ratings Regarding the Usefulness of Informal Types of Information for Learning About the Quality of Education Provided at the School (n=33)

How useful is each type of information listed below in helping you learn about the quality of education provided at this school?

	Not at all useful				Very useful	Blank/
	1	2	3	4	5	Missing
a. Report cards	12% (n=4)	18% (n=6)	27% (n=9)	12% (n=4)	24% (n=8)	6% (n=2)
b. Talking with the teacher	6% (n=2)	3% (n=1)	12% (n=4)	6% (n=2)	70% (n=23)	3% (n=1)
d. Seeing graded samples of my child's work	6% (n=2)	3% (n=1)	9% (n=3)	24% (n=8)	55% (n=18)	3% (n=1)

Table 8

Parent Interview Responses Regarding the Usefulness of Standardized Tests for Learning About the Quality of Education Provided at the School (n=33)

How useful is each type of information listed below in helping you learn about the quality of education provided at this school?

	Not at all useful				Very useful	Blank/
	1	2	3	4	5	Missing
c. Standardized tests	18% (n=6)	6% (n=2)	24% (n=8)	21% (n=7)	24% (n=8)	6% (n=2)

Normative information n=12 ($\bar{X}$ =4.2)

"The standardized test would probably be very useful because they compare with other students in other schools."

"That is one of the reasons I like the standardized tests, because to me if you have a national standard test for third graders, it shows you where your kid is against national standards which doesn't necessarily say anything about your kid but it might point out there is a problem here . . ."

"The standardized testing, I guess, if it's compared to other school districts, it would have to be somewhat useful in understanding the quality of education."

"The standardized tests—there are so many things that fall into each district, I don't know how I feel about those. Yes, I think it would. Because of course the parents are going to want their child to go to the better schools."

Table 8 (continued)

Curriculum information n=4 ($\bar{X}$=4.0)

"It shows you what the kids are doing, what he does know, what he can remember, what he's learning in the class. It gives you an idea about what they're presenting to the kids."

"I would feel that that would be very useful with the school. That they're following certain guidelines.

Other n=9 ($\bar{X}$=2.9)

"As parents, we just go by what we have here. We don't make a comparison as far as what's going on elsewhere. We just kind of take it for granted what's here."

One parent replied that the standardized tests aren't useful because her child has learning disabilities and has a hard time reading the test.

Child hasn't taken test or parents haven't seen results n=3 ($\bar{X}$=2.3)

Test not important n=2 ($\bar{X}$=2.0)

"I just don't feel it's that important."

Test provides no information n=2 ($\bar{X}$=1.0)

"I don't think that shows any effect on the school even."

"That doesn't tell me anything about the quality of education here at this particular school. It just tells me that these are tests that pretty much had to be given out during the end quarter of the semester."

Unfamiliar with standardized tests n=1 (parents left question blank)

When rating informal sources of information, the results from the interviews can be seen as a more extreme version of the pattern found in the questionnaire data (as occurred for the first set of questions on children's progress). Even for the purpose of evaluating the school, parents find talking to the teacher and seeing graded samples of work to be more useful than standardized tests. In Table 9 categories of parent responses are presented with illustrative quotations. In support of their positive ratings, parents explained that talking with the teacher allowed them to see what expectations were being set, the quality of the curriculum, and how caring the teacher is with students. Specifically, parents said that seeing the actual work that students brought home let them judge whether what was being taught was worthwhile.

Table 9

Parent Interview Responses Regarding the Usefulness of Informal Types of Information for Learning About the Quality of Education Provided at the School (n=33)

Talking with the teacher helps identify her expectations and goals (n=4)

"Talking with them gives you a better feel of the goals that they have in mind for the kids, the criteria, teacher criteria, the goals for different grade levels. Usually when we talk about their progress it's in relationship to those kind of things which is an indication to me of the quality of the education here."

Talking with the teacher is telling about her personality and style with students (n=7)

"I think if there's a lot of kids that are having problems, that maybe it reflects the teacher. You know, so to me, that's where I would say how it affects the school."

"I think that's very useful, very significant to tell you that the teachers are really caring, or that they're really good with the students."

"Talking with the teacher tells me a great deal about the quality of education because she is able to tell me about herself, you know, her strengths, her weaknesses, and things like that."

Seeing graded samples (and to a lesser extent, talking with teacher) allows an evaluation of the curriculum (n=15)

"That [graded samples] lets me know whether my child's learning and I think the paperwork that they get lets me know what they're doing."

"I think [graded samples are] very useful. I like to know what kind of things because not everything that they teach, you think is useful."

"I just think if it's a really good cross section of what they do, the samples, then it would be a pretty good indication of the quality. If it's just random cases of their work, or ones that they just save because they like and they put them in the portfolio, or something, then I don't think it's a real good indication."

"I can see what kind of work the teacher is handing out. The teacher is the one that's in there quarterbacking the classroom. You know, if she's handing out pretty basic stuff to the kids to work on, then that's pretty boring, you know, "get-me-through-the-school-day" type of activities as far as I'm concerned. But if she's handing out stuff that will keep their interest and get their initiative going as far as keeping them active in school, and wanting to learn, that pretty much sets the tone for the school year and gives me an indication of what kind of quality teachers there are, and what kind of quality programs are here at this school."

Report cards are not useful because they're about the child (n=3)

"That's a 2 because that doesn't tell anything about the school, that tells about the student."

Report cards are not useful because of subjective grading (n=4)

"[They] wouldn't be very useful in telling the quality of education because you're not really sure what they are being graded on until it has been explained to you or you have seen other things."

". . . because everybody grades differently."

Report cards are not useful—Other reasons (n=10)

Note. Data from the three questions in Table 7 were treated as an aggregate. The themes emerged from parent comments. Categories are not mutually exclusive.

> I can see what kind of work the teacher is handing out. The teacher is the one that's in there quarter-backing the classroom. You know, if she's handing out pretty basic stuff to the kids to work on, then that's pretty boring, you know, "get-me-through-the-school-day" type of activities as far as I'm concerned. But if she's handing out stuff that will keep their interest and get their initiative going as far as keeping them active in school, and wanting to learn, that pretty much sets the tone for the school year and gives me an indication of what kind of quality teachers there are, and what kind of quality programs are here at this school.

The only consistently negative comments regarding informal measures pertained to report cards. About half of the sample said that report cards were not very useful for judging the school because "they don't tell about the school" or because grading criteria are ambiguous or noncomparable across schools.

Parent Reactions to Standardized Tests and Performance Assessments: Questionnaire Results

On the questionnaire, parents were provided with displays showing typical multiple-choice questions that appear on traditional, standardized achievement tests and a sample of more open-ended questions used in performance assessments. Separate displays were shown for both reading and math. Although more examples were given on the questionnaire, the following are illustrative of each type:

Standardized test questions	**Performance assessment problems**
How much change will you get if you have $6.55 and spend $4.32? ( ) $2.23 ( ) $2.43 ( ) $3.23 ( ) $10.87	Suppose you couldn't remember what 8 x 7 is. How could you figure it out?
A good title for this story would be— ( ) The Pet Store ( ) A Strange Occupation ( ) Going on a Trip ( ) Danny's Pet Snake	At the end of the story, the little old man and the little old woman go on living in the little old house. Do you think this is a good ending? Tell why or why not.

Parents were then asked to rate their approval or disapproval of each type of measure. Respondents were not forced to choose one type of measure over the other. In fact, as can be seen from the data in Table 10, many parents liked both. (As a consequence, oblique factor analyses showed no correlation between the "standardized testing" factor and the "performance assessment" factor.) Although the majority of parents approved of both types of measures, performance assessments had higher approval ratings than did standardized tests, especially in the "Strongly approve" category. For example, in mathematics 18% strongly approved of standardized tests whereas 31% strongly approved of the use of performance assessments.

To try to get at the issue of the influence of assessment formats on instruction, parents were also presented with the following choices:

> Currently in the nation there is some concern that standardized tests have had an undue influence on what gets taught in school, because of too much "teaching to the test." **IF** it is true that tests influence instruction, indicate below whether you prefer for children to be taught mathematics and reading using questions like Box A or Box B:

Box A contained the sample standardized test questions for each subject and Box B contained the performance assessment problems.

As shown in Table 10, more parents preferred the use of questions like those on performance assessments for instructional purposes than preferred the use of standardized-test type questions. Although the preference for more open-ended questions is quite large in reading (58% favored performance assessments compared to 21% liking standardized tests), the difference was narrower for instruction in math (44% favored performance assessments over 31% for standardized tests). The tendency for a subgroup of parents to differentiate their choices by subject area is consistent with a pattern found later in the interview data and can clearly be connected in more extended interview responses to parents' conception of mathematics as a one-right-answer subject area. A different, small subgroup of parents resisted choosing between the two types and added comments in the margin emphasizing that a combination of both should be used.

Table 10

Parent Questionnaire Approval Ratings of Standardized Tests and Performance Assessments (n=105)

Suppose the school district were trying to decide whether to use standardized tests, performance assessments, or <u>both</u> standardized tests <u>and</u> performance assessments in all its elementary schools. Indicate below the degree of your approval or disapproval of each type of test or assessment.

Do you approve or disapprove of:	Strongly disapprove	Disapprove	Neutral	Approve	Strongly approve
12. using standardized tests with questions like those in **Box A** to measure your child's achievement <u>in math</u>?	2%	10%	24%	46%	18%
13. using performance assessments with questions like those in **Box B** to measure your child's achievement <u>in math</u>?	6%	5%	20%	38%	31%

Do you approve or disapprove of:	Strongly disapprove	Disapprove	Neutral	Approve	Strongly approve
14. using standardized tests with questions like those in **Box A** to measure your child's achievement <u>in reading</u>?	2%	11%	22%	47%	18%
15. using performance assessments with questions like those in **Box B** to measure your child's achievement <u>in reading</u>?	2%	7%	15%	47%	30%

Currently in the nation there is some concern that standardized tests have had an undue influence on what gets taught in school, because of too much "teaching to the test." <u>IF</u> it is true that tests influence instruction, indicate below whether you would prefer for children to be taught mathematics and reading using questions like Box A or Box B:

	Strongly prefer instruction like Box A	Prefer instruction like Box A	Neutral	Prefer instruction like Box B	Strongly prefer instruction like Box B	Blank/ Missing
16. For math instruction	15%	16%	21%	27%	17%	4%
17. For reading instruction	9%	12%	17%	37%	21%	4%

Once again we wished to examine whether the lowered enthusiasm for standardized tests in this sample compared to the nation could explain the more positive response to performance assessments. Therefore, responses to the above questionnaire items were again considered separately for the subgroups that "Favor" and "Oppose" standardized national tests on the Gallup Poll. Not surprisingly, the group that favors standardized tests gave higher approval ratings to standardized test questions than did the Oppose group (Math: $\overline{X}$ = 4.10, 2.97 respectively; Reading: $\overline{X}$ = 4.13, 2.87 respectively). However, as suggested above by zero correlations between the two factors, approval of standardized tests does not imply disapproval of performance assessments. Parents who favor standardized national tests rated performance assessments just as highly as did parents who oppose standardized national tests (Math: $\overline{X}$ = 3.92, 3.61 respectively; Reading: $\overline{X}$ = 3.90, 3.87 respectively). Yet, when forced to choose between the two types of questions for use in classroom instruction, there were significant differences between the two groups. Parents who favor standardized tests rated the two types of measures about equally in the context of math instruction ($\overline{X}$ = 2.72) whereas parents who oppose standardized national tests expressed a stronger preference for the use of performance assessments in mathematics instruction ($\overline{X}$ = 3.80). In reading, the pattern was similar ($\overline{X}$ = 3.17, 3.93 respectively). Thus population shifts in the proportion endorsing standardized testing do affect the degree of approval of performance assessments in some contexts, but these analyses suggest that even if 100% of the national population answered yes to the need for standardized national tests, that answer would signify approximately equal approval of performance assessments, not disapproval of them.

Parent Reactions to Standardized Tests and Performance Assessments: Interview Results

It was the impression of the interviewers that almost all parents were intrigued by the opportunity as part of the interview to have a close look at both standardized test questions and performance assessments for third graders. Despite being presented with these examples near the end of the interview, most parents took time to look through the materials carefully. Many worked through the sample problems and asked questions about how they were administered typically or whether this was third-grade work. As can be seen from sample transcript data in the Appendix, parents often pointed to the two different types

of questions or used specific problems as examples when explaining their preferences for standardized tests or performance assessments.

Because the purpose of the interview was to understand the reasoning behind parent preferences, qualitative analysis was used to develop categories representing different positions. Entire transcript segments from this portion of the interview were read and distilled so that shortened segments could be used that still contained all key ideas from the longer transcripts. Each representative segment was given a shortened label; similar segments were sorted into categories resulting in the final categorization scheme shown in Table 11. In addition, a separate coding system was developed for key ideas because it was discovered that many features of the two types of tests, cited repeatedly by parents in their discussions, bore little relation to parents' final recommendations or preferences about which type of measure should be used. For example, parents in both the standardized testing and performance assessment camps commented on the worry that performance assessments might be unfair because of the extra demand they placed on students to write their answers.

Table 11

Parent Interview Evaluations of Standardized Test Questions and Performance Assessments: Categories of Responses (n=33)

Strongly prefers standardized tests (n = 2)
Prefers standardized tests (Likes both) (n = 1)
Standardized tests Math/ Performance assessment Reading (n = 4)
Standardized tests Reading/ Performance assessment Math (n = 1)
Both: Likes both, Wants both in instruction (n = 2)
Prefers performance assessments (Likes both) (n = 11)
Strongly prefers performance assessments (n = 11)

The categories in Table 11 are ordered roughly from the position most favorable toward standardized tests to that most favorable toward performance assessments. Only 3 of the 33 parents or parent dyads preferred the use of standardized tests for both district and instructional purposes. Respondents in this category saw standardized tests as more cut and dried, more aligned to instruction, and easier (because having the answers there made it clear what was expected). Another subgroup of parents preferred standardized tests for math and performance assessments for reading. This group of parents liked the opportunity for children to express themselves and the possibility for more than one right answer for reading assessment but maintained that there is one right answer for math. Two sample transcript segments are shown for this category in the Appendix. Like them, a third respondent explained: "In mathematics, I think this type [ST] is probably the best . . . because math is pretty basic as far as having the right answer and you have to have the right answer. With this [PA reading] they can use their imagination and they can tell you a story the way they see it. It doesn't always have to be one way." One respondent preferred standardized tests for reading and performance assessments for math. Two responses were assigned in the "middle" category, liking both types of measures for both district purposes and for classroom instruction.

By far the majority of respondents preferred performance assessments. Twelve interview segments were placed in the category "Prefers performance assessments (Likes both)" and another 11 responses were in the "Strongly prefers performance assessments" category. Although the "Prefers PA (Likes both)" category was heterogeneous, the dominant response was to approve of both kinds of measures being used for district purposes but to prefer that performance assessments be used for classroom instruction. Across responses in all categories the most frequently mentioned feature of performance assessments is that they make children think. As shown by the counts for each code in Table 12, this was noted by 10 of the 12 responses in the "Prefers PA" category. In addition to the several responses shown in the Appendix, other respondents in this category explained why they preferred performance assessments for instruction:

Table 12

Key Features of Standardized Tests and Performance Assessments Identified in Interviews by Subgroups of Parents

Codes	Strongly prefers ST (n=2)	Prefers ST (Likes both) (n=1)	ST Math/ PA Reading (n=4)	ST Reading/ PA Math (n=1)	Likes both (n=2)	Prefers PA (Likes both) (n=12)	Strongly prefers PA (n=11)
Guess				1		6	9
Y/N	2	1	3			5	2
Easy	1		1	1	1	2	2
Obj	1					2	2
Support		1				3	1
Know (ST)					1	1	1
Norm						2	1
Think		1	2	1	2	10	8
Imagin	1		4			5	4
Diag	1		1			8	6
Hard	1	1		1	1	5	1
Unfair	1		1		1	5	2
Subj	1				1	3	3
Know (PA)					1	4	4
Align		1	1			4	3

Note. The following codes were used in the transcript analysis for frequently mentioned characteristics of standardized tests and performance assessments:

Guess	ST allows students to guess
Y/N	ST have black and white, yes-no answers
Easy	ST are easy
Obj	ST are objective, have a clear right answer
Support	Having the answers on ST helps children know what's expected
Know (ST)	ST shows that children have skills, can comprehend
Norm	ST provides national comparison
Think	PA makes students think
Imagin	PA lets children express themselves, show imagination, be creative
Diag	PA shows you where children need help
Hard	PA are difficult, + a challenge, – too hard
Unfair	PA reading and writing demands might be unfair to some children
Subj	PA are difficult to grade, subjective
Know (PA)	PA shows children really have the concepts because they have to explain
Align	Would have to teach more like (PA) to test that way

> I think this [PA] would make them think more.
>
> I think in order to learn any kind of subject you have to have concepts down and I think number 2 is going to show how to develop the concepts better. . . . You need to get those basics. . . . But I do think this [PA] is going to make them think more.
>
> I just think, like here [ST] they get to choose where here [PA] they really have to think about it and you're really gonna know that they know exactly what they read or what have you.

Even respondents who preferred standardized tests for other reasons often noted that performance assessments would stimulate children's imagination or make them have to think.

The substance of comments regarding both standardized tests and performance assessments was similar for parents in both the "Prefers PA" and "Strongly prefers PA" categories with the distinction tending to be that parents in the first group gave approval ratings to standardized tests for district purposes. Both groups, however, criticized standardized tests because it was possible to guess and because the results of such tests would not tell you much about the process of a student's learning or where a student was failing to comprehend. In contrast, we used the code "Diag" to note the many instances where parents commented on the diagnostic value of performance assessments. For example, from one of the "Prefers PA" responses not shown in the Appendix:

> The other tests [PA math] kind of makes them tell you the concept, not just the right answer. I like the "explain your choice," or "what would you tell Adam" type questions. . . . this would give a teacher more information to think about, especially on the concepts that they haven't quite grasped yet.

Several parents also commented that standardized tests were more objective in contrast to performance tests that would be either difficult to grade fairly or more time consuming to grade. In some cases, however, parents went on to say that this was exactly the kind of attention or commitment that was needed in education. Concerns about grading were voiced along with the concern that reading level might create problems for some students on the math assessment or that writing might be a problem on both performance assessments. Although these concerns were mentioned as problems to be

resolved, for the most part they did not appear to affect parents' enthusiasm for using performance assessments in instruction. For some parents in the "Prefers PA (Likes both)" category, however, these features were cited specifically as the reason that both types of tests should be used at least for district purposes. "Different children learn in different ways." Some children, especially those who "are not good with words," would be helped by having the answers there so they could show that they understood. (Answers of this type were coded "Support" in Table 12.)

Conclusions

The purpose of the study was to examine parent opinions about standardized tests and new performance assessments in greater depth than can be understood from national survey data. The classic Gallup Poll question showing a high percentage of citizens and public school parents in favor of standardized national tests is often interpreted as a mandate for external, machine-scorable accountability measures. What was discovered in this study is that parents' favorable ratings of standardized national tests do not imply a preference for such measures over other less formal sources of information for monitoring their child's academic progress *or* for judging the quality of education provided at their local school. Approval of standardized tests likewise does not imply disapproval of performance assessments.

In this study, third-grade parents considered report cards, hearing from the teacher about their child's progress, and seeing graded samples of student work to be much more useful in learning about their child's progress than standardized tests. In interview data, parents often mentioned the need for comparative information to know how to interpret their own child's progress, but they trusted the teacher to tell them how their child was doing in relation to grade-level expectations or to other children in the class. Comparison to external or national norms was mentioned rarely. Even for accountability purposes, the usefulness ratings for standardized tests increased but did not equal parents' high ratings of talking to the teacher and seeing student work. According to parents, these sources of information are indicators of school quality because they allow parents to see what is being taught and what expectations are set by the classroom teacher.

Because preliminary analyses of Gallup Poll results showed the third-grade parent sample in this study to be less favorably disposed toward standardized tests than the national sample, follow-up analyses were conducted to determine whether high parent ratings of informal sources of information could be attributed to the particular nature of this sample. However, even the 46% subsample of parents who favored standardized national tests on the Gallup Poll rated report cards, talking to the teacher, and seeing student work as more useful that standardized tests.

Using both questionnaire and interview formats, different samples of parents were provided with specific examples of the types of questions used on standardized tests and on performance assessments in both reading and mathematics. While a majority of parents approve of both standardized tests and performance assessments, approval ratings were stronger for performance assessments. Again this overall pattern could not be explained as merely a selection bias in this particular sample. The group that favored national tests rated the two types of measures equally. A pervasive theme in the interview data was that performance assessment problems "make children think" and are likely to give teachers better insights about what children are understanding and where they are struggling. Parents commented frequently about the desirability of having children explain their answers in mathematics and being encouraged to express themselves in response to stories they read. Standardized tests were seen as easier and more supportive by some parents because having answer choices communicates what's expected and allows children who aren't very verbal to show what they know; at the same time, parents complained frequently that multiple-choice questions allow children to guess the right answer "25% of the time."

When parents in this study had a chance to look closely at performance assessment problems, most endorsed their use for district purposes and especially preferred their use in classroom contexts. Therefore, survey data showing approval of standardized national tests should not be taken to mean that parents are opposed to the use of alternative measures. A few parents expressed concern about the subjectivity of performance measures, but lack of rigor—sometimes evoked by the term "alternative assessment"—was never mentioned. Instead, when parents looked at specific performance assessment problems, they were more likely to comment that the questions seemed hard or

challenging. It was the impression of the interview team, in fact, that seeing real examples of both types of tests affected parent responses. The questionnaire and interview demonstrations served an educative function. Parents might have answered differently if they had to rely on their own ideas about what performance assessments look like.

References

Borko, H., Flory, M., & Cumbo, K. (1993, April). *Teachers' ideas and practices about assessment and instruction.* Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.

Darling-Hammond, L. (1988). Accountability and teacher professionalism. *American Educator*, *12*, 8-13, 38-43.

Elam, S.M., Rose, L.C., & Gallup, A.M. (1992, September). The 24th annual Gallup-Phi Delta Kappan Poll of the public's attitudes toward the public schools. *Phi Delta Kappan*, *74*, 41-53.

Flexer, R.J., & Gerstner, E.A. (1993, April). *Dilemmas and issues for teachers developing performance assessments in mathematics.* Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.

Frederiksen, J.R., & Collins, A. (1989). A systems approach to educational testing. *Educational Researcher*, *18*(9), 27-32.

Hiebert, E.H., & Davinroy, K. (1993, April). *Dilemmas and issues in implementing classroom-based assessments for literacy.* Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.

Madaus, G.S., West, M.M., Harmon, M.C., Lomax, R.G., & Viator, K.A. (1992). *The influence of testing on teaching math and science in grades 4-12: Executive summary.* Boston: Boston College, Center for the Study of Testing.

McNeil, L.M. (1988). Contradictions of control, Part 3: Contradictions of reform. *Phi Delta Kappan*, *69*, 478-485.

Resnick, L.B., & Resnick, D.P. (1992). Assessing the thinking curriculum: New tools for educational reform. In B.R. Gifford & M.C. O'Connor (Eds.), *Changing assessments: Alternative views of aptitude, achievement and instruction* (pp. 37-75). Boston: Kluwer Academic Publishers.

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

Smith, M.L. (1991). Put to the test: The effects of external testing on teachers. *Educational Researcher*, *20*(5), 8-11.

Wiggins, G. (1989). Teaching to the (authentic) test. *Educational Leadership*, *46*, 41-46.

Appendix

Parent Interview Evaluations of Standardized Test Questions and Performance Assessments: Categories of Response With Sample Excerpts (n = 33)

Note: Parents were shown samples of standardized test questions and performance assessments. They were asked whether they approved or disapproved of each type. Then they were asked to discuss the strengths and weaknesses of each. Half of the interviews presented the math comparison first, and half presented the reading comparison first. The last question asked parents what type of questions they would prefer to see used in classroom instruction.

The standardized test prompt and performance assessment prompt are abbreviated ST and PA below. Parent ratings of strongly approve are shown as ++, strongly disapprove as – –, and so on. The question about instruction is shown by the word (Instruction) in parentheses. When a mother and father both participated in the interview their responses are denoted M and F.

STRONGLY PREFERS STANDARDIZED TESTS (n = 2)

06

(STreading++) Well, I think it's really clear cut what is expected of these kids. It's easy for them to understand, it's easy for them to answer it.

(PAreading–) Well, the disadvantages of it are that there are too many right or wrong answers. I think that is kind of hard for kids that age to comprehend all this. Maybe the advantages of it would be they are more able to use their imagination.

(PAmath–) For one thing, if the child might understand how to do something like this but they don't know how to explain it. They have trouble with words . . .

Just that a test like this might be useful again to get an idea how they are at comprehending different things but it wouldn't really be fair to grade their learning on this.

(Instruction) Oh. Standardized. Because I feel that it is easier to teach and easier for the kids to learn and easier to grade them on it.

PREFERS STANDARDIZED TESTS (LIKES BOTH) (n = 1)

01

(STreading++) When I take tests, I like to take this kind.

(PAreading+) Well, maybe advantage-wise, it would make 'em pay closer attention.

(STmath++) I think, when you're taking a test, that you're nervous anyway, and if you have the right answer somewhere in there [laughs], you're gonna do better . . .

Well, this could be the other side of that coin. Sometimes the disadvantage is knowing that the right answer is there somewhere.

(PA+) . . . Yeah, I don't really like that "explain your choice." I think that they're doing pretty good to pick out the right one without explaining why they did it.

(Instruction) . . . If they're being taught this way and they're doing, you know, they seem to be doing well, in this kind of test (PA), then I have no problems with that. If they're being taught to do it.

STANDARDIZED TESTS MATH/PERFORMANCE ASSESSMENTS READING (n = 4)

33

(STmath+) I would rather the standard achievement. It's more clear what they want from the child than the performance one. It's more clear and more understandable.

(PAreading+) On this one [reading] I would have to choose the performance, yeah. Because it gives the child a little more chance to express his own personal views and this one [ST] there's just the choices.

(Instruction) I'd say the performance assessments [in reading] 'cause it still does give him a chance to still tell their part. [In math?] I would have to say I prefer the standardized 'cause that's not like an option, there's only one answer, you know.

11

(PAreading++) Strongly approve. I like that because they have to think. It's not as cut and dried. Here [standardized] you just "a," "b," "c," or "d" . . . and this one [performance] is more . . . There's a little more leeway, individuality.

(PAmath–) 'Cause if you're measuring math, I mean I expect things like this on the reading, but on a lot of these there's a lot of . . . If you have a kid who's a math whiz and can't read a lick, they're gonna flunk something like this. Basically, they wouldn't understand what they have to do if they cannot read. Here [standardized] it's pretty simple; there's a plus [+] there's a minus [–].

(Instruction) These [performance], all of them. [In both subject areas, then?] Yeah. 'Cause it makes 'em think. And again, thinking about using these for math, no, I don't think is right because they're learning how to do everything else. But if they had these kind of questions in school, they'll be more familiar.

STANDARDIZED TESTS READING/PERFORMANCE ASSESSMENTS MATH (n = 1)

08

(STreading+) I like this one better. I think they would get lost by the end of the story [PA]. At her grade level, I don't know if she would remember all these things. Without going back and re-reading. . . . Well this one [ST] was like they're getting the understanding of the words, and this all in one short thing so their mind still has all this fresh in their mind.

(PAreading–) . . . Probably for kids who are really, really far ahead in their class . . . it might be good for them.

(STmath,neutral) It doesn't explain enough to keep their mind going. I think they would probably just guess at all the answers.

(PAmath++) Yeah. 'Cause it explains to them more how to think it out.

(Instruction) I think they should think up their own things. 'Cause all kids are different. The teacher should think of a broader horizon of things and not just go by the book.

BOTH: LIKES BOTH, WANTS BOTH IN INSTRUCTION (n = 2)

17

(STreading+) (PAreading+) I think that's good [PA] because you have to really—If they have to write it in paragraph form then that means they have really read it well and understood it. . . . I feel like they need to understand the story in more detail than the standardized tests. But one of the bad things about it is the kids who aren't good at putting their thoughts into words are going to have problems with it also.

(STmath+) I think that these standardized tests would be easier for the kids to understand.

(PAmath+) These would go like according to their grade level and everything, right? I would approve. It makes them really think.

(Instruction) In reading? I suppose this one because it does make them, you know, if they learn from the beginning they really have to think, they really have to comprehend. This will probably teach them better. [In mathematics?] Um, that's a hard one. I feel like this one has always been fine. I guess I never thought about doing it different because this has been done for so many years. This [PA] would be more of a variety of teaching probably. [If you had to pick one?] I think this one. [ST]

PREFERS PERFORMANCE ASSESSMENTS (LIKES BOTH) (n = 11)

14

M: [You should have two different kinds of tests] because there's children that think one way and there's children that think another way.

F: I think both should be used and I think both should be used in the same test. I mean, one's subjective and one's an objective evaluation. . . . I believe the advantages of multiple-choice type questions are that a student that has had a hard time expressing themselves is much, is [more] able to show that they understand what they have read. . . . I believe a disadvantage of the multiple choice is that children can guess. . . . Maybe their achievement isn't actually going to be shown as well [with ST] as it might be on this [PA], as you have to write a summary, or write the answers to the questions.

(PAreading+) F: That is, for a child that they can express themselves through writing. This is probably a much better way of evaluating them than the multiple choice, but I believe it does have possible disadvantages for some children.

(Instruction) M: These kinds of questions [PA]. F: I agree. M: Yeah, because it makes them think . . .

F: I believe that this gives—for most children—a more true evaluation and it would be harder to teach the test.

M: There's more of an understanding why it's right and wrong, you know, with this way. And the kids, I think, should be taught in math and reading and stuff like that, they should be taught to understand the rights and wrongs of it.

39

(STmath+) The standardized test would measure what is being taught, more typical of what they're working on in their math books. My older child is getting more of [PA] in 6th grade, emphasizes thinking.

(PAmath–) I don't necessarily think that it measures the way that they're taught to solve problems. . . . This is certainly going to make your child think more . . . but it would be extremely difficult if they can't read.

(STreading+) (PAreading++) This [PA] looks like a more fun test. Is it just to measure reading, would be doing quite a bit of writing? More interesting test, pictures would capture their imagination. They could look at [ST] and just guess. I would choose [PA] over [ST].

(Instruction) [They should use] questions that make children think, not just a yes or no. . . . In math there is a right answer, but still [our teacher] has them write down the steps about how did we get it. That kind of thing is really good in math, something I didn't learn in math, they have to look at each step and explain. You should have questions that make them analyze their answers.

28

(STmath+) (PAmath+) F: One of the disadvantages, you know, on the multiple choice, you could just put an "x" there and have a chance of getting it right. Where over here on the other ones you have to explain why you made that choice, so you have to think your answer through.

M: It might be a little overwhelming to some kids. . . . But yet it might would help the teacher by saying well, they're not quite comprehending everything, and see where they're struggling, you know.

(STreading–) (PAreading+) M: Well, I feel basically the same way as he does. You know, I think that this [ST] is more the way that things were. The other one is better. It makes the child have to think more . . . and they have to kind of use more of their imagination, too.

15

(STreading+) F: I would like to see both of them used. You are going to be pretty much right or wrong on it. . . . M: where there is one right answer.

(PAreading++) . . . F: The creativity gives a chance to come to their own conclusion. And there is . . . not everything that is so black and white where it has to be yes or no. M: Every child is going to have a different view. It lets them express why they thought it was that way.

(STmath+) F: There's enough situations through life, where you have to have an exact, it's good to know the correct way. M: I think, I don't know, that a disadvantage might be—It's kind of funny but in my college classes I used to call it "multiple guess" instead of multiple choice.

(PAmath++) M: . . . I think that it gives the child a chance to show their work and how they came up with the conclusion of the answer . . . and the teacher can see, well, if the child filled this out and she could pinpoint something and say, well, this is why you are not coming up with the right answer.

(Instruction) F: I would go more for the ones like these [PA]. I feel it would keep more excitement in the class and I think it would keep kids interested. M: . . . the kids have to think more. It's a little bit harder. F: they [could] walk away and not remember what they checked in the box whereas something like this [PA] they are going to think about and stimulate them more.

20

(STreading+) (PAreading++) . . . Actually, I'd like to see them have this type of test but you have to start really young with them, showing them how to communicate and how to really write that out, bring it out of themselves. . . . I would go with this one [PA]. I would strongly approve of this type. I want my son to learn how to write more, communicate better. . . . This seems the faster way [ST] as far as a test time goes, but this [PA] looks like they've really worked out the problem. They've had to sit there and think about it and take the time to do it.

(STmath+) I'm comfortable with these still, so I approve of them.

(PAmath+) This would be interesting. I'd like to see them start working some of these into the program.

(Instruction) I would like to see them use these ones [PA], because as I'm looking at this, you're reading it and it's asking you, and it's almost as though you're talking to the teacher one on one. As you're looking at this [ST], you say 4 times 8, what is that? Well, this one [PA] is giving you a little bit more challenge. It's kind of almost speaking, you say OK, now you figure this out. "Suppose you couldn't remember what 8 x 7 is. How could you figure it out?" It seems like this is better communicated this way.

22

(STmath++) (PAmath++) This [ST] shows exactly where a student is. . . . This makes the student do a little more thinking. They don't see the answer right away. They have to figure it out.

(STreading+) (PAreading++) This [PA] helps with memory skills. They have to read it. They have to go back over it to answer the questions at the end. It helps with comprehension. They must come up with their own answers. No answer is wrong. I have trouble comprehending what I read. If I had had to do things like this, it would have helped my comprehension.

(Instruction) (PA) They focus on the whole spectrum. They don't have to comprehend with the standardized reading test. They would have to read the whole thing with this [PA]. This would be better at judging where the student is at. It pertains more to actual life situations. The answers are not in front of them. They have to think about it.

STRONGLY PREFERS PERFORMANCE ASSESSMENTS (n = 11)

31

(STreading+) F: The main thing I see about this is that it is a non-judgmental way to compare the kids. It gives you, maybe not a perfect handle but it gives you comparison on how kids are doing in the school system which may point out flaws in the school systems or flaws in the kids. As far as the teacher goes, because it is standardized that you know exactly what the right answer is. . . . The only disadvantage of this is, is that I have always in my life been able to pass any multiple choice whether I have an inkling of what the answer or not.

(PAreading++) M: I see advantages because he can express himself. . . . To me, you are getting to know more about your child through this one than you would on the other [ST]. F: Exactly. You get a lot more handle on whether or not he actually comprehended what he read if he has to go ahead and rephrase what he's read and give a synopsis. . . . The disadvantage to this test [PA], the big one is the grading is so subjective. The grading of it depends totally on the quality of the teacher or the person sitting down to grade it because it is totally subjective.

(STmath–) F: This to me, by handing them the answers again doesn't really test their skills. Somebody that doesn't even grasp addition could guess on these answers. M: I just think these [ST] are good for comparison just like the other ones but I like these [PA] better.

(PAmath++) F: . . . I love these because they actually make the kid read the problem and come up with the answer instead of picking it from a list that is right here beside it. . . . M: The disadvantage I can see is, if your kid is not up to reading level for this, he is going to be much better on the other one.

(Instruction) F: For classroom instruction, this [PA] is the only way to go. In the classroom to sit down and say, which of these is right? That doesn't make them come up with the solution. M: I feel to make my kids well rounded, he needs to express himself. These [ST] don't do it. To me, I hate rote.

05

(STreading–) . . . M: It doesn't really force you to think, I mean, the answers are right there. . . . F: It makes you have one, their choices instead of one of your own choices.

(PAreading+) . . . M: This one has you also like explain. Like right here [ST] you don't really have to think too much about it, and this one [PA] you really have to kind of pull it all together—and reason it out. . . . M: The only problem I see with this if they were at a lower level of third-grade reading, you know, they probably couldn't grasp some of this.

(STmath–) (PAmath++) F: If I had the option, something like this [PA] would be a little bit better. . . . M: Yeah. I think this would make them, if they were to teach, obviously they'd have to teach this to take these tests. We'd probably get better quality in teaching. Things that would probably stick with them a little bit more. . . . M: I think you can probably guess more on [ST]. On these [PA] you can't really guess, you kind of have to think about it. F: Plus, I think this [PA] makes it a little bit more interesting for the kids. This is pretty cut and dried.

(Instruction) F: Well, like we said, this one. This one, I think the kids could relate to this. . . . M: It's more practical. You can apply to it, like it kind of stresses more of life skills.

16

(STreading+) (PAreading+) This one [PA] will tell you more about what they get out of the book. . . . or the story. Just by the questions they ask, and that you have to re-write.

(STmath–) Well I just don't like the pick or choose. (PAmath+) I approve of this. They have to work out the problems.

(Instruction) This [PA]. They make them work on 'em more . . . the other ones they could [just] choose.

19

(STreading–) (PAreading++) F: Okay. This test [ST] is not very subjective. You know, there's a right or a wrong answer, but it doesn't require or doesn't allow for a lot of input. . . . I think it does test their comprehension because . . . there're some good questions there. You know, "What is a good title for the story?" I think is a good question, 'cause then they're gonna have to understand and comprehend the overall picture. . . . This test [PA], on the other hand, does require them to write something out and . . . M: It also causes them to think for themselves a little bit more. F: Yeah, the answer's not provided for them, they have to fill in the answer. . . . F: I just think it'd be tougher to grade this test. . . . M: But this test would say a lot more, it

would tell a lot more. F: The advantage to this [PA] is I can't guess. On the other one [ST] I can guess and I have a 25% chance of getting the right answer.

(STmath–) (PAmath+) . . . M: This one [ST], it gives you an example of what is right, it gives you one answer that is right, so you can come up with the right answer fairly easily. F: It's not gonna help them. M: It's not gonna help you to understand the solution to the problem a number of different ways like this one [PA]. . . . F: This one [PA], by writing out the answer in most of these on the other test, they're going to show you the process that they used to come up with the answer, so I think it goes beyond what this test [ST] asks, which is for the right answer.

(Instruction) M: Well . . . I like the fact that [. . .] these tests [PA] here help you to help the child think about what their answer is. F: And I think it'd be important to ask questions that help the child understand how they're solving problems and how they're comprehending things. . . . F: I don't think that this standardized achievement test allows the child to understand why he, how he comes up with the right answer. I think the performance assessment does, you know, it allows them to actually give examples of how they solve the problem.

04

(STreading–) Well, to me these are very simple to grade . . . I mean, very simplistic. I mean, it's right or wrong. You know, it's an easy man's way out. [laughs] It requires little imagination.

(PAreading++) I would strongly approve of the way to test [my daughter's] reading skills and things. . . . It asks her a great, a variety, it asks her to use her mind to remember, I mean, she has to pay attention to detail. She has to evaluate her own, she has to learn to make a judgment call.

(Instruction) I definitely like the one that allows them to express themselves. . . . I do. I think that it's, it takes more time on the part of the educator, but I think that the education system would benefit immensely from it. I think you pick up, you know, learning deficiencies much quicker with a performance assessment than you would with standardized. I think that they have a good chance of picking out the [right answer] just by even guessing. . . . But with having to explain it, you see their abilities, you see whether they can or cannot explain. . . . It shows their understanding level, where this one [ST] doesn't.

36

(STmath–) (PAmath+) I prefer this [PA], because no matter what, they can still guess on this [ST] and possibly get it right. On this they're going to have to get it right. . . . They're going to have to figure out how to come up with the answer.

(STreading– (PAreading+) . . . This way [PA] you have to read it, they have to give a reason, they have to solve the problem. This you have to tell why, you know, why did it have a good ending? The whole thing. You have to explain yourself.

(Instruction) Well like this [PA]. Ones that they have to explain—tell the answer, and then tell why they think that.