

Effects of High-Stakes Testing on Instruction

Lorrie A. Shepard
Katharine Cutts Dougherty
University of Colorado at Boulder

Traditionally standardized tests of achievement were used to report to parents and to monitor state and district trends. Such tests had almost no effect on instruction because teachers paid so little attention to them. According to a national survey conducted by Goslin in 1967, teachers only infrequently used the results of standardized tests and reported virtually no influence of test content on teaching methods or course content. This picture changed, however, beginning with the

bad news for proponents of measurement-driven instruction. Student performance has improved, according to the NAEP, in basic reading and math skills during the period of high-stakes testing. At the same time, however, there has been no gain or a decrease in higher-order, advanced skills. Many attribute this pattern of results to the negative influence of standardized testing on teaching and learning.

Evidence documenting the negative influence of testing on instruction is limited to a few studies. Darling-Hammond and Wise (1985) reported that teachers in their study were pressured to "teach the test." Teachers gave specific examples

both standardized tests and alternative tests, to see if students really knew what they

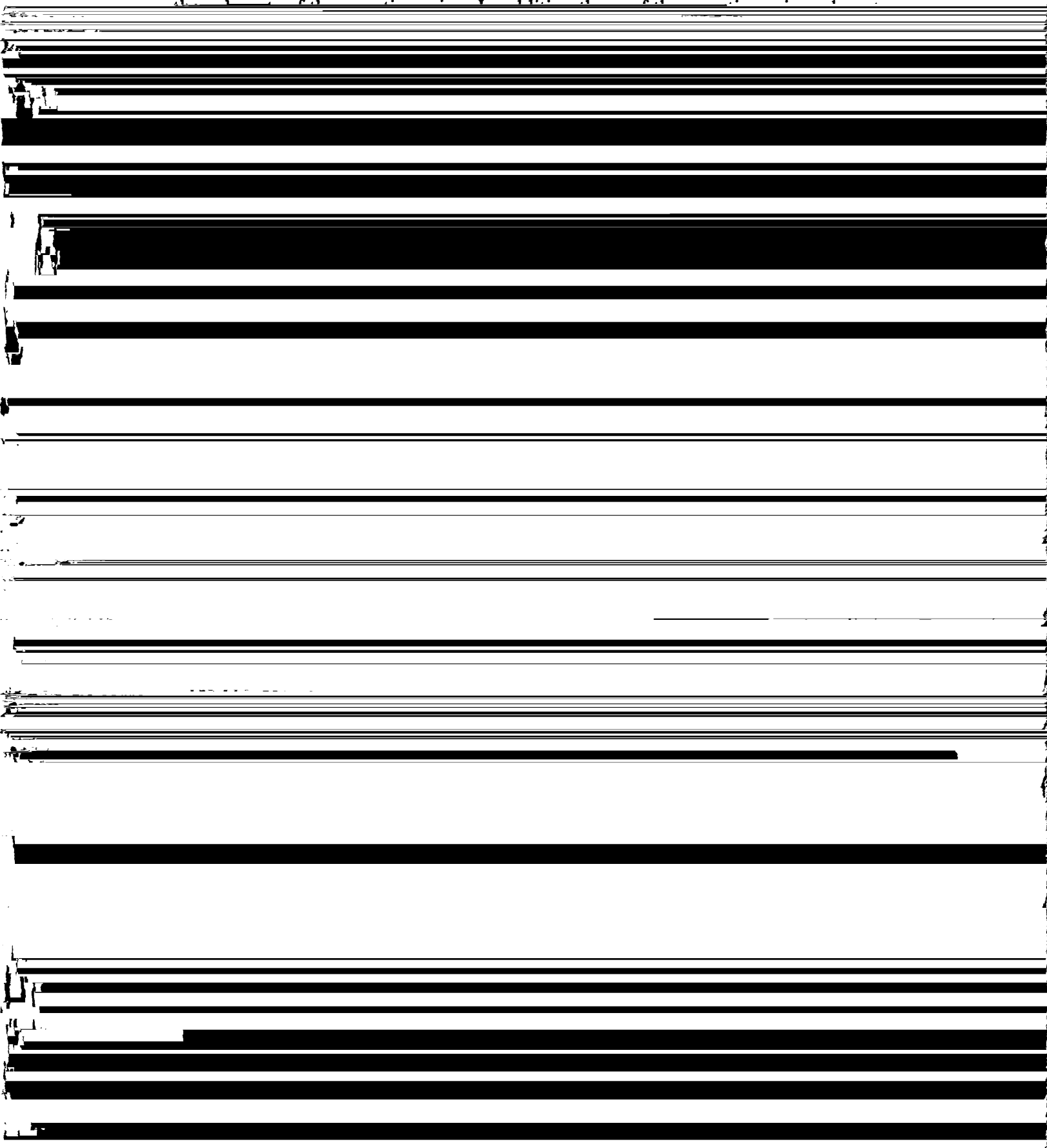
~~they had learned on their public account high school test. The test~~

... 11 ... 3 ... until after the districts' normal standardized

testing period in April. It was also decided that the teacher questionnaire should not be administered until after all standardized testing had been completed for the year. As a result the teacher questionnaires were distributed the third week in May, 1990 with only two weeks remaining until the end of the school year. This timing undoubtedly reduced return rates.

To calculate response rates we estimated the numbers of teachers in the

Whereas in advance we had anticipated that there might be one general factor explaining attitudes toward tests, pressure for accountability, and instructional effects, instead we found that there were discrete factors corresponding to each of



Even within these districts where testing was thought to be pervasively high-

... associated with other effects. For

[The remainder of the page is heavily obscured by horizontal black bars, likely representing redacted text or scanning artifacts.]

thinking activities" and "practice in divergent problem solving," 63% and 57% respectively.

When the results in Tables 3 and 4 are taken together it is clear that the great majority of teachers give greater attention to basic skills because of standardized

~~tests. Some teachers follow the predictable pattern of increasing practice on the~~

In addition to the previously mentioned correlation between test preparation and instructional emphasis on basic skills, test preparation was also correlated with ~~the use of multiple-choice questions (r = .26)~~ and with the use of multiple

monetary district resources." Therefore, teachers may not have accurate information about all of the uses of test scores.

Half of the teachers reported that tests are used occasionally or frequently "to compare or evaluate teachers;" 45% said that they are occasionally or frequently used to evaluate principals; and 53% said tests are used occasionally or frequently "to determine awards for school excellence." On the first three items, evaluating teachers, evaluating principals, and allocating district funds there were significant differences between the two districts, with District B reporting these uses more often.

In general external uses of test results were reported to be much more frequent. Tests are used to compare districts, to rank schools in the newspaper, and to reward schools; 76%, 76%, and 71% of teachers said these uses occurred

Four of the Measurement-Driven Instruction items were agreed to by a majority of teachers: "I spend more time teaching reading and math and less time teaching social studies and science because reading and math test scores are so important" (64%), "I use fill-in-the-blank worksheets and matching exercises in my regular instruction so that my students will be comfortable with short answer formats when it comes time to test" (60%), "When I teach reading and math, I emphasize the skills and content I know are on the standardized tests" (69%), and "A lot of the workbook and textbook activities I select for students to do are very similar to the short passages and stand-alone questions that students will encounter on tests" (62%).

Test-driven items with which the majority of teachers disagreed were: "higher-order thinking skills are something I get to only if there is time after covering the basics" (61% disagreed), and "I don't use essay tests during the year because I

Time spent giving standardized tests

Two final questions were asked about the number of hours spent giving

students." We did not add in the duplicate counts for teachers' second and third answers because it would make it difficult for the reader to make the numbers "add

" However, it should be remembered that many of the high frequency answers

reminded to practice practice for the test. The fun and excitement has been taken

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

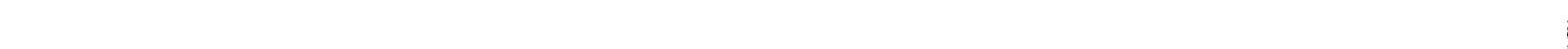
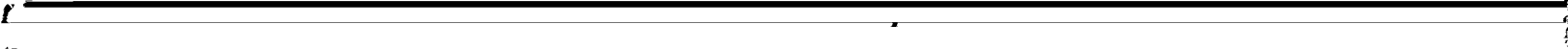
[REDACTED]

[REDACTED]

[REDACTED]

3. Because of the importance of standardized tests, teachers give greater

~~the link between the report of controversial testing practices and measurement driven~~



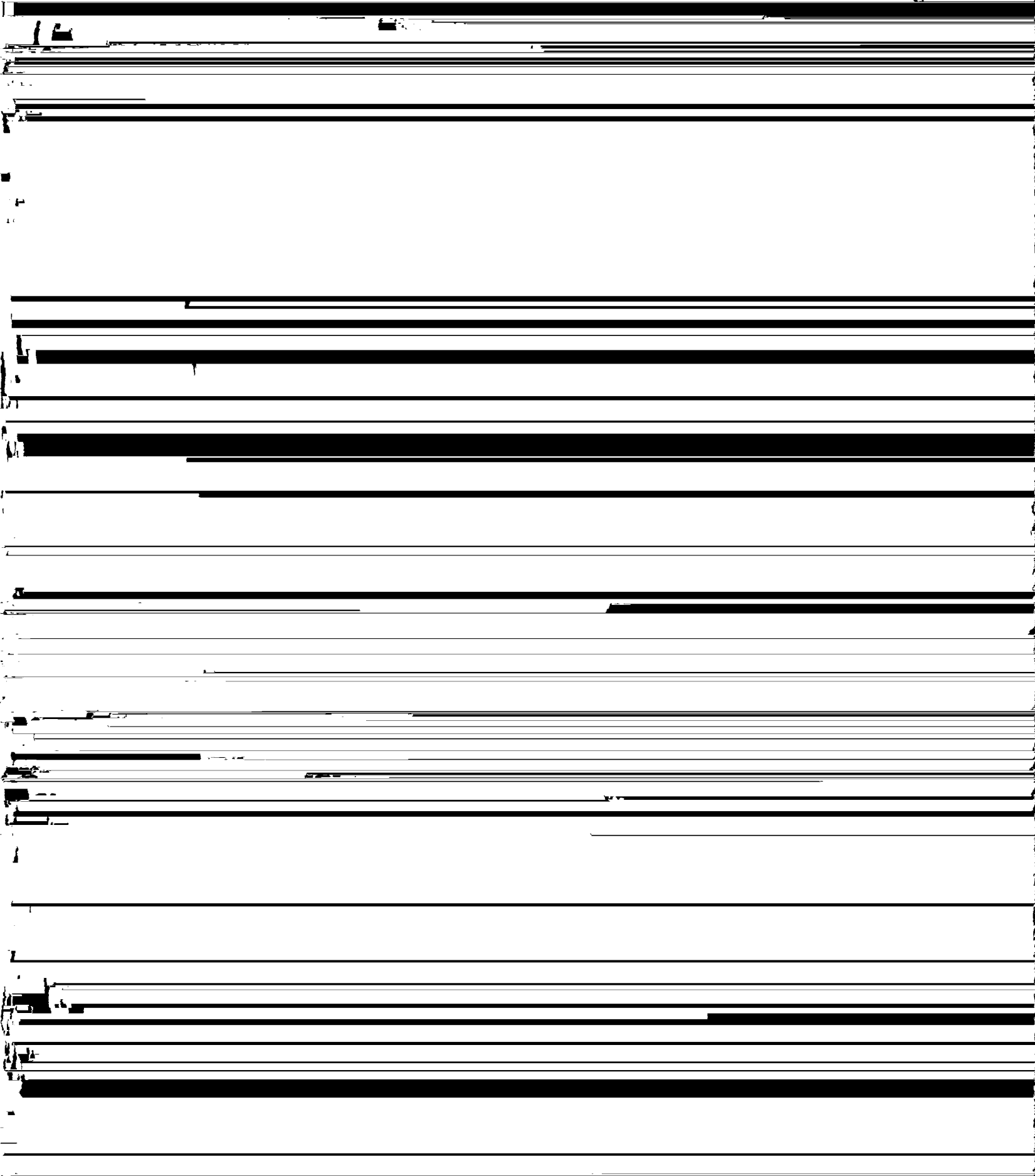


Table 1

Factor Analysis of the Teacher Questionnaire

Oblique Rotation Pattern Matrix with Nine Factors Extracted (N=120-350)

	FACTOR 1 (contro)	FACTOR 2 (skillins)	FACTOR 3 (testprep)	FACTOR 4 (extuse)	FACTOR 5 (pro-tests)	FACTOR 6 (measdriv)	FACTOR 7 (divergen)	FACTOR 8 (pressure)	FACTOR 9 (intuse)
CONTR38	.87	.02	-.04	.01	-.05	.04	.09	-.03	.01
CONTR45	.84	.04	-.06	.02	-.05	.05	.10	-.05	.02
CONTR36	.81	.03	.02	.03	.11	-.01	-.07	-.03	-.04
CONTR41	.80	.02	.01	.01	.07	.03	.08	.00	.02
CONTR47	.80	.05	.01	.01	.00	.01	.01	.00	.05
CONTR39	.77	.03	.10	.06	-.10	-.30	-.20	-.15	.05
CONTR40	.72	.10	-.09	-.09	.00	.03	-.07	.05	.08
CONTR41	.70	.08	.04	.13	-.06	-.07	-.03	-.02	.04
CONTR42	.65	-.06	-.02	-.16	-.02	-.04	-.08	.08	-.04
CONTR46	.58	-.04	.03	.09	.09	-.14	-.03	.07	-.09
CONTR43	.57	.01	.00	-.02	.06	-.01	-.03	.09	-.09
USE47	.26	-.02	-.12	.10	-.08	.03	.06	-.04	-.22
INST9	.10	.85	.09	.06	.04	-.01	-.03	.00	.10
INST10	-.06	.83	.04	.05	.00	.00	-.04	.04	.09
INST11	-.01	.73	.00	.02	.02	-.05	.08	-.02	.01
INST13	.11	.67	-.04	.01	.01	.08	.10	.05	.02
INST19	-.02	.46	-.02	.05	-.05	-.25	.05	.03	-.03
INST12	.07	.44	-.13	-.03	.10	.03	.07	.00	-.06
INST20	-.06	.30	-.19	-.19	.04	-.06	-.01	.01	-.08
PREP30	.00	-.05	-.87	.13	-.06	-.05	.07	.05	.10
PREP31	.00	.00	.01	.01	.00	.00	.00	.00	.00

Table 2

ITEM MEANS AND FREQUENCIES AND SUBTEST STATISTICS
FOR TWO HIGH-STAKES DISTRICTS ON THE SUBSCALE: PRESSURE

PRESSURE FOR IMPROVED TEST SCORES. To what extent do you feel pressure from the following groups to improve your students' standardized test scores?

No Pressure	Slight Pressure	Moderate Pressure	Substantial Pressure	Great Pressure	<i>Place a check in the appropriate column to show how much pressure to improve test scores you personally feel from each of the groups.</i>
-------------	-----------------	-------------------	----------------------	----------------	--

Table 3

ITEM MEANS AND FREQUENCIES AND SUBTEST STATISTICS FOR
TWO HIGH-STAKES DISTRICTS ON THE SUBSCALE: SKILLS INSTRUCTION (SKILLINS)

TESTING AND INSTRUCTION. Items 6-28 describe a variety of instructional activities. Circle the appropriate response to show whether the use of standardized tests in your school causes you to place MORE EMPHASIS or LESS EMPHASIS on the activity than if there were no mandated standardized tests, or whether your emphasis on

- | | | | | |
|-----|----------------------------------|---------------------|----------------------|---------------|
| 9. | Basic skills in mathematics..... | More Emphasis | Not Influenced | Less Emphasis |
| 10. | Basic skills in reading..... | More Emphasis | Not Influenced | Less Emphasis |
| 11. | Reading for understanding..... | More Emphasis | Not Influenced | Less Emphasis |

Table 4

ITEM MEANS AND FREQUENCIES AND SUBTEST STATISTICS FOR TWO HIGH-STAKES

6. Subjects which are not tested.....	More Emphasis	Not Influenced	Less Emphasis
14. Kids talking about what's been read.....	More Emphasis	Not Influenced	Less Emphasis
15. Extended project work in mathematics.....	More Emphasis	Not Influenced	Less Emphasis

Table 5

ITEM MEANS AND FREQUENCIES AND SUBTEST STATISTICS FOR
TWO HIGH-STAKES DISTRICTS ON THE SUBSCALE: TEST PREPARATION (TESTPREP)

TEST PREPARATION ACTIVITIES. For items 29-35, try to recall how much time over this entire school year you

spent on the following test preparation activities. Check one box for each item.

29. Give students worksheets that review the content you expect to be on the test

Table 6

ITEM MEANS AND FREQUENCIES AND SURTEST STATISTICS FOR TWO HIGH-STAKES

DISTRICTS ON THE SUBSCALE: CONTROVERSIAL TESTING PRACTICES (CONTRV)

students' test scores. To what extent do you believe they are practiced by teachers in your school? Use this 5-point scale to tell us the extent to which each of the practices listed in items 36 to 46 occurs in your school; simply circle the code that matches the frequency of each practice.

Table 7

ITEM MEANS AND FREQUENCIES AND SUBTEST STATISTICS FOR TWO HIGH-STAKES DISTRICTS ON THE SUBSCALE, INTERNAL USES OF TESTS (INTUCE)

[The table content is heavily obscured by horizontal black bars and is therefore illegible.]

Table 8

ITEM MEANS AND FREQUENCIES AND SUBTEST STATISTICS FOR
TWO HIGH-STAKES DISTRICTS ON THE SUBSCALE: EXTERNAL USES OF TESTS (EXTUSE)

USES OF TEST DATA. Items 47-57 list some uses of standardized test scores. Use the same 5-point scale to indicate how frequently standardized test scores are used in your district for the following purposes.

50. To compare district schools against one another.....N.....R.....O.....F.....?
51. To compare district performance against other districts.....N.....R.....O.....F.....?
52. To make decisions about what curriculum to emphasize.....N.....R.....O.....F.....?
-
53. To decide on continuation of innovative programs.....N.....R.....O.....F.....?
54. To rank schools in the newspaper.....N.....R.....O.....F.....?

Frequencies in %					
#	N (1)	R (2)	O (3)	F (4)	? a
50	3.0	3.9	13.6	71.2	8.3
51	1.9	3.0	12.7	75.9	6.4

Means		p<.001
District A	District B	
3.55	3.70	
3.61	3.77	

Table 9

ITEM MEANS AND FREQUENCIES AND SUBTEST STATISTICS FOR TWO HIGH-STAKES
DISTRICTS ON THE SURSCALE PRO STANDARDIZED TESTING (PRO TESTS)

THE EFFECTS OF STANDARDIZED TESTING. Today in the United States there is intense debate about whether
the standardized testing imposed on schools for accountability purposes is good or bad. A number of

11

potential positive and negative effects of standardized tests are listed in items 50-77

Table 11

TEAM MEANS AND FREQUENCIES AND OTHER STATISTICS FOR THE NINE TEAMS

Team	Mean	Frequency	Other Statistics
1			
2			
3			
4			
5			
6			
7			
8			
9			

Table 12

FREQUENCIES FOR HOURS AND DAYS SPENT GIVING STANDARDIZED TESTS

80. Estimate how much time you have spent in your classroom this school year giving standardized tests.
(Administering the reading and math subtests of a typical standardized test takes about 4-5 hours.)

- Less than 4 hours 4-8 hours 9-16 hours 17 or more hours

81. How many days have been interrupted this school year as a result of giving standardized tests?

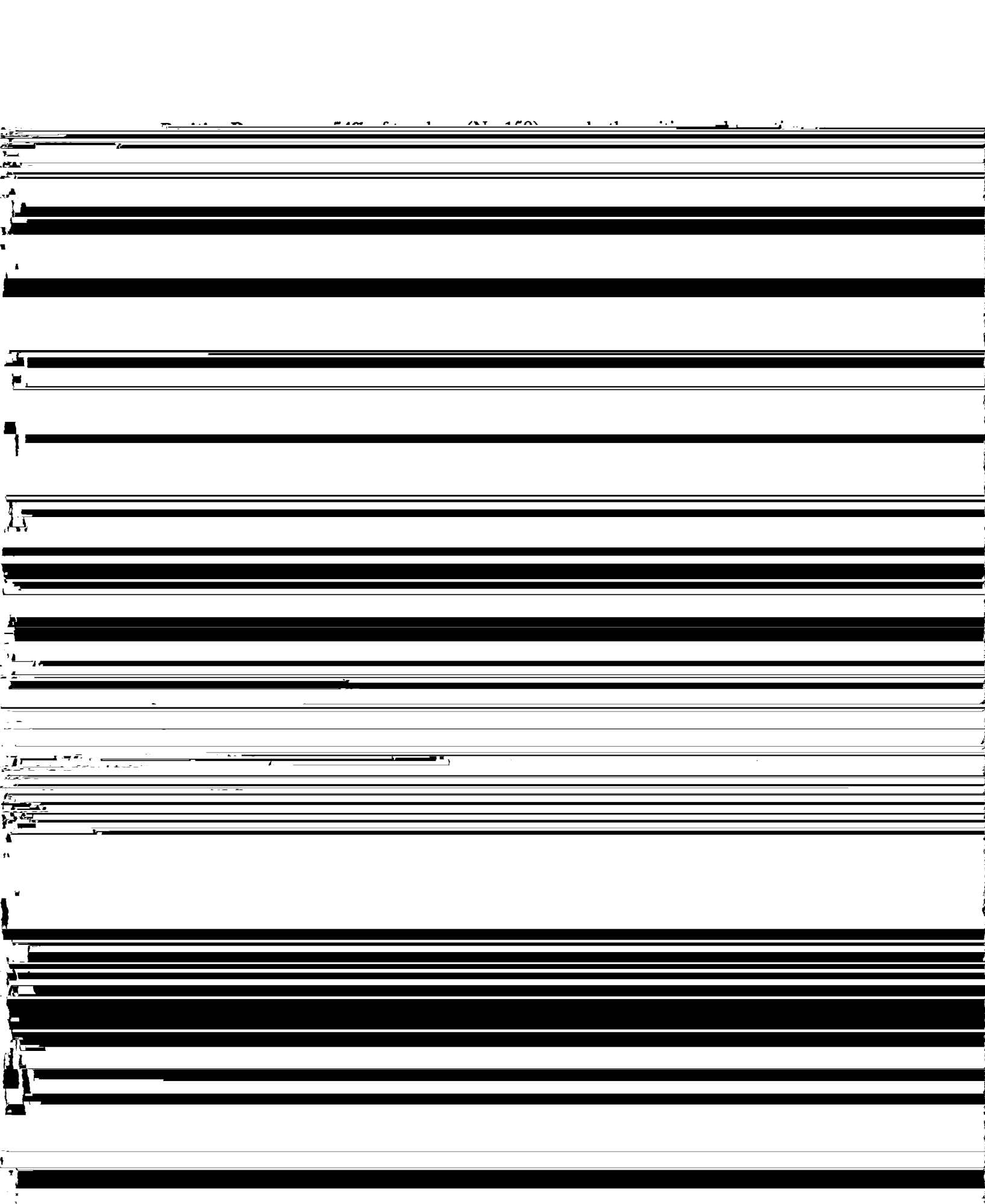
Table 13
CORRELATIONS AMONG SUBSCALES AND ODD ITEMS

	Testprep	Controv	Intuse	Extuse	Protests	Measdriv	Essays	Multip	Timecom	Promretn	Testhrs	Intdays
	.87											
	.06	.94										
	.03	.42	.84									
	.00	.24	.61	.84								
	.12	-.16	-.08	-.09	.81							
	.26	.45	.21	.21	.10	.82						
	.42	-.08	.13	-.02	.14	.04	.87					
	.26	.09	.02	-.03	.01	.10	.06	1.00				
	.18	.05	.03	.06	.02	.14	.00	.09	1.00			
	.19	.37	.44	.26	-.11	.14	.07	.08	.05	1.00		
	.22	.19	.00	-.04	-.13	.15	.04	.01	-.05	.12	1.00	
	.21	.20	.16	.08	-.18	.20	.14	-.02	-.08	.14	.47	1.00

stency coefficients are reported on the diagonal. Ns for
wise comparisons ranged from 102 (for the MEASDRIV by
lation) to 346.
gnificant at $p < .01$ are in bold face type.

Table 14
Summary of Teachers' Answers to Open-ended Questions
on the Positive and Negative Effects of Standardized Testing
(District B, N = 280)

~~Original Dimensions with Standardized Tests. The debate about testing pro and~~



"After I have given pretests, I can analyze the results and concentrate on what areas are low."

"Identifying weakness, specifically in comprehension, has caused us to emphasize critical thinking skills and allowed greater flexibility in exploring new techniques."

Identify strengths and weakness of individual students for extra help: (N=39)

"Saw weaknesses in pretest and individualized instruction or used small group instruction to help master basics."

"When the test was used as a diagnostic tool to determine strengths and weaknesses."

"In some cases such tests have revealed weaknesses in time to remediate such weaknesses."

"It gave me a chance to know what to focus on if a child was weak in a particular area."

"Standardized tests do provide a starting point for emphasizing students' strengths and weaknesses. The tests also lead parents and teachers to discussions of individual students that can be helpful."

"In looking at test scores I am able to help student with the area that they are having the most trouble with."

"It helps to see what categories they are have mastered and those that are

"By taking the standardized tests the students felt like they were working toward a common goal. A hurdle to overcome. Getting high scores on the practice tests made them feel smarter."

Used in placement and grouping decisions: (N=2)

"Standardized tests are part of the prerequisite to get into the gifted class and advanced classes."

"Preparing my reading and math groups."

Provide norm-referenced information: (N=2)

"It helps to know where students rank among the nation."

Other positive responses: (N=13)

e.g., "We team taught for (the state test) and found it good for students and

"It enables pupils to do critical thinking."

"They are simple straight to the point directions."

"Affirm current classroom results."

Negative Responses: 54% of teachers (N=150) gave both positive and negative

"Too much time is needed to emphasize test content, test taking skills, practice work-sheets."

"Limits creative teaching--You are always teaching (or so it seems) towards the

~~test. You are always teaching towards the test.~~

"Because the test results are 'flaunted' in the news etc., I become very resentful at having to give the tests. The pressure is bad on teachers and therefore passed on to students. It's as if school's out when tests are completed."

"Timed for mastery; some students require more time than most tests allocate.
The grading process is biased in reference to some students' environment."

Only Negative: 24% (N=66) gave only negative answers, some said specifically that they could not think of positive answers to question 78. Categories of only negative responses are as follows:

Too much teaching to test content and test format: (N=28)

"Repetitive, boring drill. Use of district materials required too much time."

"Standardized test has had a negative influence on my teaching because some enrichment as well as practical skills must be set aside."

"I feel very sorry for the children in (State) whose school administrators and media cause them to miss out on so many areas of learning. Many teachers

believe the children are suffering from it."

Negative or unfair comparisons: (N=7)

"Our standardized test scores are typed up and passed around to every teacher

to let them know how they are doing. We also pass them to the teachers with low