The Texas Teacher Test

LORIE A. SHEPARD AMELIA E. KREITZER

The TECAT, the Texas Examination of Current Administrators and Teachers, is a test of basic literacy that was given to Texas teachers in March, 1986. The test, seen as politically essential to leverage a tax increase and pay raise for teachers, was intended to raise the public esteem of teachers by weeding out incompetents. Teachers expended massive effort in reviewing basic skills and drilling on test format. After two tries, 99% of the 216,000 who took the test had passed. Shop teachers, special education teachers, and coaches were overrepresented among the failures. The costs of district-sponsored workshops and the incentive day to take the test brought its public cost to a sum 10 times greater than policy makers had anticipated. Though most teachers agreed that literacy skills are prerequisite to good teaching, paradoxically, most also reported that being threatened by a low-level test of fundamental skills was demoralizing. Ironically, many think that the TECAT damaged public esteem for teachers because stories about incompetence in teaching and portrayals of teachers' qualifications appeared alongside examples from a very easy test.

Texas is one of three states with legislated mandates to test the competence of practicing educators. In March of 1986, 202,000 teachers and school administrators took the TECAT, the Texas Examination of Current Administrators and Teachers, to see if they could keep their jobs. Though direct TECAT results were not used for focused research because it was a highly visible example of a reform strategy contemplated by other states, policymakers who wish to improve public education see several options. Increasing high school graduation requirements, lengthening the school day, testing students, providing preschool education, changing funding formulae, and so forth. Testing teachers particularly appeals to reform-minded decision-makers because it is a concrete and decisive action aimed directly at the quality of education in the classroom.

The Texas teacher test was also deemed worthy of systematic investigation because it even generated a new genre of tests accompanying the educational reform movement. Unlike large-scale assessments used in the past to measure the effects of program change, current reformers now use tests themselves as the instruments of change. Tests of this type are potentially powerful interventions intended directly to raise standards either by spurring examiners to improve skills or by removing examiners with inadequate performance. These reforming tests, akin to what Popham (1986) calls high-stakes tests, are not well understood by either measurement specialists or policy makers. Laws to create these tests were passed based on beliefs about their effects; opponents have different beliefs but no more compelling evidence.

This article summarizes a case study (Shepard, Kreitzer, & Graue, 1987) aimed at understanding both the context and the effects of the Texas test. Can lessons from the Texas experience inform policy decisions in other states? What were the educational problems and political context that gave rise to the reform legislatures? What did advocates and opponents believe the effects of testing would be? What were the effects of the test? Who failed? What was known about the teaching competence of those who failed? What can be said about the impact of the teacher test on the quality of education and public confidence in schools? How much did the testing program cost and were the benefits worth the cost?

Research Methods

The study involved an amalgam of social science research methods. Key political figures and informants were identified from newspaper accounts and by asking each respondent for the name and phone number of central participants. Structured interviews were conducted face-to-face with individuals centrally involved in the creation of educational reform legislation and its implementation. The chairman of the House Public Education Committee and bill author, chairman of the Senate Education Committee, Speaker of the House, legislative aides, the Commissioner of Education, presidents and lobbyists for the four teacher organizations, and Texas Education Agency (TEA) directors of assessment. Respondents were asked to describe the political climate and key events that led to the inclusion of a teacher test in the reform legislation; to characterize their own role in the enactment of the legislation; to depict the positions of proponents and opponents; to relate their perceptions of the impact of TECAT; and to give advice to legislators in other states.

Interview sessions were audio recorded, transcribed, and subsequently analyzed in two stages. Transcript segments were coded as answers to prequestion data sets or as emerging themes or issues. Identifying labels were assigned to new codes to link recurring themes across interviews. For example, "no test, no tax" was a code used to tag references to the political bargain (ultimatum) made in the last days before the passage of the legislation. In later reiterations of the
transcripts, subtypes or competing positions were identified. Exemplars were collected of each position and theme.

Stories and editorials from two Texas newspapers were collected systematically. From the Austin American Statesman, articles were assembled under the following headings: "The TECAT un-til April 1986," "'The 1984 special legis-lative session,'" and "'The Select Com- mittee on Public Education.'" From the archives of the Amarillo News-Globe, articles were obtained for three files: "School testing," "'1984 special legis-lative session,'" and "H. Ross Perot." The same descriptors used in these files guided the collection of a small set of articles from the New York Times. Clippings were also obtained unsys- tematically from seven major Texas newspapers. Prior to the first site visit in May of 1986, newspaper accounts were used to construct a chronology of previous reforms and events leading to the administration of TECAT. Later, a content analysis was undertaken of the two complete newspaper files to deter- mine how public school teachers were portrayed by the press.

Original documents were collected and examined. The Texas Education Agency provided copies of TECAT descriptive materials and the TEA- produced Study Guide, as well as background memos and data. For ex- ample, the Commissioner's letter to the State Board was available with the data used to facilitate standard setting. The test contractors had written profes- sional papers describing test develop- ment procedures. The Governor's office granted access to the files of the Select Committee on Public Education, which included data such as SAT scores, transcripts of testimony from educator groups, and drafts of findings from the subcommittees. Interview re- sponses were often supplied relevant documents from their files, including transcripts of speeches, instructional materials, and newsletters. If infor- mants referred to data or reports, sources were retrieved. The Public Policy Research Laboratory, Texas A & M University, which conducts a Texas public opinion poll, was con- tacted for a complete set of releases on education issues. A representative sample of 100 Texas teachers and 50 administrators were selected using a two-stage sampling strategy. First, school districts were stratified by size and a stratified-random sample of 20 districts was selected. Then, educators were ran-domly selected to represent the straturn proportions. These teachers and administrators were interviewed by telephone, following a structured protocol, during the summer after the TECAT administration. The response rate was 96%. Based on the demo- graphic characteristics of the non-respondents, the reported results could reflect a slightly positive bias. The interviews were recorded, transcribed, and analyzed following the procedures described for the key figure interviews. In the case of the teacher interviews, several of the ques- tions could also be summarized quanti- tatively, as in the number of hours spent studying for the test.

The authors attended a total of five days of workshop sessions offered by regional Service Centers to help teachers prepare for the test. The sites were selected to achieve geographic and demographic spread. Of the several sites using University of Texas developed video tapes for instruction, only one was chosen. Researchers were divided access to three sites because of the amount of stress involved for the participants who were retaking the test. The preparation sessions were recorded, transcribed, and analyzed. Instructors were interviewed. Work- shop participants completed teacher questionnaires on a voluntary basis. Although responses could not be con- sidered representative, the questionnaires provided answers and writing samples from teachers who had failed TECAT the first time.

A representative sample of personel directors was interviewed in the fall of 1986, by which time teachers were either to have passed TECAT or been fired. In telephone interviews using a standard protocol, directors were asked about the general impact of TECAT, about the teaching qualifications of teachers who failed, and about the inci- dence of teachers not taking the test. Other questions pertained to the specific treatment of teachers who failed and to the more general handling of teacher shortages in that district. Methods of analysis described pre- viously were again applied to written transcripts of these interviews.

The research methods in this study can be described as both preordinate and responsive (Stake, 1975). Many of the study populations and research questions were determined before- hand. Other aspects of the study were developed in response to issues and questions identified in the early stages of data collection. For example, once it was known that Education Service Centers played a major role in pro- viding preparation and review for the test, personnel at all 20 centers were contacted and asked standard ques- tions about the instruction provided and the region's population of failed examinees. Copies of instructional materials were requested and schedules were obtained for remaining work- shops. Phone calls were made to cable TV stations in several major cities to document the extensive schedules of TECAT broadcasts which had been described in teacher interviews. The methods of the field work are described in greater detail in the full-length technical report (Sherpa et al., 1987).

Political Context

Until 1980, oil had held Texas immune from economic troubles felt by the rest of the country. Texas had gone through the 1970s without a state in- come tax and without raising the sales tax, all on the profits from oil. But when the world-wide energy glut reduced the price of crude oil, the state's dependency on oil revenues turned the boom to bust overnight. Participants recalled that political rhetoric about correcting the ills of public education was tied to business interests and accelerated when Texas first felt the effects of economic recession.

In 1982 Governor Mark White had been elected after promising to seek a 24% pay increase for teachers, but failed to get the necessary tax increase from the legislature. In 1983 several events came together to create the impetus for major reform with an accompanying tax increase. A Nation at Risk and other national reports added to the dismay in Texas about the state of public education. When we asked aides, legis- lators, and teacher leaders to think back to specific evidence of educational decline, they recalled that: "Texas was near the bottom compared to other states on SAT scores" and "Across the nation the weakest college graduates appeared to be going into teaching." Also in 1983, failing to obtain a budget increase, the governor ap-
pointed a Select Committee on Public Education charged by multi-millionaire, H. Ross Perot. Many say it was the power and visibility of Perot that gave this blue-ribbon panel clout not enjoyed by previous advisory committees. He had a penchant for oneliners that kept the Committee's work in the news for a year (e.g., "scheduling academic subjects around band and sports is joke."). He reportedly spent one-half million dollars to effect the work of the Select Committee and hired his own lobbyists to see reforms through the Special Session of the legislature. Business leaders heard from one of their own that Perot could not hope to compete for high technology investments if northern executives were unwilling to move their families to Texas schools.

The Select Committee heard testimony about the need to upgrade the profession of teaching, to pay higher salaries to attract the best people into teaching, and to keep the most talented from leaving. In the process they collected horrifying stories about the incompetence of some practicing teachers. Apparently one teacher was said to have had difficulty explaining to her class why the weather was so different in Hawaii and in Alaska even though they were right next to each other (in the corner of the map). A formal survey revealed that most educators believed that up to 10% of their colleagues were incompetent. Furthermore, a significant percentage of superintendents, principals, and teachers felt that it was difficult to fire a bad teacher (Sirota & Alper, 1984).

Recruiting news stories emphasized that a substantial number of teachers lacked basic skills. In Houston, 62% of new teachers failed the Pre-Professional Skills Test; in Dallas, the superintendents explained that he was hiring below standard minority teachers to satisfy a court desegregation ruling, and an equal number of unqualified white teachers to avoid reverse discrimination (Dallas Times Herald, 12/13/85). Inflammatory committee members shared stories about lessons they had received from teachers arguing for pay raises that were peppered with bad grammar and esthetic values.

Teacher testing to weed out incompetent teachers emerged from the Select Committee as a necessary element in educational reform. A test for practicing teachers might have been inevitable from that point on, although the advisability of testing was hotly debated in the legislative Special Session. The inevitability of a teacher test was sealed when it became a bargaining chip to leverage a tax increase. Lobbyists and aides heard over and over again the phrase, "no test, no tax," attributed to Stan Schluter, Chairman of the House Ways and Means Committee. A tax bill would not get through his committee without evidence that higher wages would not go to incompetent teachers. As many legislators voiced it, "We have to be accountable to our constituents. We are not going to pass any kind of tax bill if we can't assure them that there aren't teachers who can't spell or read or write." The accounts we heard of the quid pro quo bargain in Texas resembled very closely rhetoric that accompanied the advent of the Arkansas teacher test.

The Texas reform legislation was an onerous bill. It provided for redistribution of resources to poor school districts, higher starting salaries for teachers, a career ladder, statewide textbook adoption, a high school graduation test, the famous "no pass, no play," rule, and a dozen more changes. Among the provisions were two levels of testing for current educators: subject-matter tests and a test of each examinee's ability to read and write. The 1984 Special Session also delivered a $4 billion dollar tax increase, the first state increase in 13 years.

The TEA CAT Day-to-day implementation decisions further shaped the Texas test. Originally, some key legislators had assumed that many teachers would be exempt from taking a basic skills test because they had already taken standardized tests such as the SAT. ETS withdrew its tests, however, because they had not been validated for the proposed use; nationally known measurement expert, James Popham, warned that even with new validity studies, the state would be vulnerable to lawsuits if teachers did not have equal access to different tests (meeting transcript, 2/18/85). To defend the job relevance of subject-matter tests, many, many more tests would be required. "An eighth grade math teacher could not be given the same test as a teacher of biology." Although the Texas Education Agency was in the process of developing ap- proximately 30 subject-matter tests for new teachers seeking certification, legal counsel advised the State Board of Education that these tests would not be defensible for practicing teachers. The board reported to the legislature that 15 million dollars would be required to implement the subject-matter tests. Commissioner Kirby estimated that a basic skills test alone would weed out 80 or 90% of incompetent teachers (American Stateman, 2/28/85). The state board proposed and the legislature accepted a basic communication test as a good faith implementation of the testing requirement; 6.5 million dollars were appropriated to develop this test and appraisal procedures for the new career ladder.

A new test. The Texas Examination of Current Administrators and Teachers, was developed to assess the minimum reading and writing skills "that practicing educators need to perform adequately in their jobs." The reading test comprised of 55 multiple-choice items that measured recognition of details and comprehension of the main ideas, job-related vocabulary, ability to distinguish fact from opinion, reference usage, and inference. The writing test included both a short composition (150 words) and a multiple-choice portion. If examinees unambiguously passed or failed on the composition, their multiple-choice answers would not be considered. Examinees who incurred in a marginal essay, however, had to pass the 30-item multiple-choice portion covering mechanics, sentence formation, and English usage in order to pass the writing test. School personnel who failed the reading or writing part of the test in March of 1986 would have a chance in June to retake the portion they had failed. Subsequent retakes were permitted, but not in time to forestall being without a certificate in September 1986. A sample selection essay is shown in Figure 1.)

TEA staff and their contractors de- voted extensive effort to constructing a test that could withstand political and legal scrutiny. A statewide Advisory Committee and a Test Review Committee were established. A statewide job-relatedness survey of 4,000 educators determined which skills were included and 1,000 educators reviewed potential test items for appropriateness and reliability. The committee proposed that passing standards be set at 75% after reviewing field-test passing.
rates for each score and recommended standards from the advisory committee and several survey groups. The commissioner’s standard would have failed 12% of the field test group; however, staff estimated that the actual failure rate would be more likely by 5% because teachers would study for the real test. The board adopted the commissioner’s passing rule in January, two months before the March testing.

Massive Preparation

One of the most unexpected findings from our research was the monumental effort that went into preparing for the TECAT. As such, the test specifications were available, the Texas Classroom Teachers Association, developed a review course and a 300-page study book. They trained 130 presenters in whom instructed 89,000 teachers in one- and two-day workshops. According to Dr. Shirley H. Crook, project director, “it was a major undertaking, logically.” On some weekends 50 workshops were going on, with 20-40 teachers in attendance at each. In addition, the University of Texas, in conjunction with the Austin Independent School District, developed 12 video tapes covering TECAT skills. These were distributed throughout the state. Most of the 20 regional service centers and many school districts purchased the tapes and checked them out to teachers or used the basis for group review sessions. One superintendent kept a VCR in his home set up for teachers “night and day” while his wife fixed popcorn. The videos were acquired by public access TV stations in major cities and shown repeatedly before both the first and second TECAT administrations. For example, in Austin the tapes covering different skills were shown every half hour for 12 hours a day for 30 days preceding each testing date.

All our major teachers organizations developed materials for workshops to ensure the success of their members. (One union forbade the use of their materials by nonmembers; other were more generous.) The largest growth was the Texas State Teachers Association, estimated that 65,000 teachers attended their workshops. Nearly every school district in the state provided test preparation opportunities. In some cases, they used the Study Guide developed by the Texas Education Agency and hired English teachers to conduct inservice sessions. Some districts made arrangements with local colleges and universities for review classes. Many districts, including large districts such as Houston, Corpus Christi, and San Antonio, paid the workshop fees for their teachers. The 20 regional service centers likewise committed themselves heavily to preparing educators for the TECAT. Several developed their own materials and practice tests; some arranged traveling workshops to reach remote areas. Many regional centers hosted six-hour video programs delivered by satellite through the Texas Interactive Instructional Network.

In our analysis of transcripts from preparations sessions we classified instructional talk as content teaching, legitimate teaching to the test, and questionable teaching to the test. Content teaching included overviews of test content, substantive presentations of few of grammar, principles of good writing, and detailed explanations or examples. All practice time was classified as content teaching, whether on sample sentences or specific test items, unless a test strategy was being emphasized rather than a substantive rule. Our conception of legitimate teaching to the test follows the Test Standards of the American Psychological Association, which urge that test takes be informed of any strategies that are “unrelated to the construct” but “influence test performance” (APA, 1985, p. 27). Many of the topics covered in the University of Texas course and materials fall into this category, for example, familiarization with test format, scoring rules, advice about guessing strategies, and anxiety reduction techniques. Examples of this type of teaching to the test are given by these excerpts from workshop presentations:

- You don’t want to leave any empty spaces. There are no penalties for guessing on this test.
- The TECAT will cover only two uses of the semicolon, both of which involve compound sentences.
- Remember you only have to know it’s wrong. You don’t have to know why, you don’t have to correct it, just know that it’s wrong.

Each of the writing workshops we attended stressed being concerned with the appearance of the essay (approx...
prize letter, legible handwriting, and correct letter format) so as to make a good impression on the scorer. Although this type of preparation should not be considered unfair or inappropriate, it would be hard to argue that the substantial time spent in these activities was really teaching teachers essential basic skills. In fact, the great majority of teachers interviewed in our probability sample said that workshops had helped them by familiarizing them with the test format and sample content rather than teaching them skills. Only three of the 96 respondents said they learned something new, specific, rules regarding punctuation and the use of pronouns.

At some point legitimate teaching to the test passed over an ill-defined line and became inappropriate. For example, after explaining that the writing samples would each be graded in one or two minutes, one instructor explained that "it's better to paragraph in the wrong place than not to paragraph at all" (because at least it would look right to the grader). This and other examples of questionable teaching to the test went beyond helping the examinees "show what he knows." Instead the strategies had the effect of helping the examinees "hide his ignorance" or use the multiple-choice format to "preshad to know." Although these strategies are not illegal (they are permitted by the test), they use clearly distort what the test can claim to have measured. The most widespread example of this second type of teaching to the test in involved exploitation of the test specifications, published by the Texas Education Agency, to "psych out" the multiple-choice question tests. The TEA Study Guide explained how the wrong alternatives would be constructed for every type of question. For inference questions on the reading test Incorrect answers would be of the following types:

- Incorrect—A statement that is controdicted by information in the reading selections;
- Unsupported—A statement that may sound reasonable, but does not necessarily follow from information in the reading selections;
- Irrelevant—A statement that is in no way logically true based on information in the reading selections. This inference often introduces information not included in the selection.

In sessions we attended, teachers practiced identifying irrelevant and specif-icly contradicted answers, so as to arrive at a correct choice among alternatives by a process of elimination. We came away thinking that teachers who were really struggling with inference would now be able to pass the items but would be unable still to recognize the valid differences from the passage.

Similar strategies for ruling out wrong answers were encouraged for main idea, detail, fact, and opinion, and even vocabulary items. The University of Texas tapes included the following information from the test author: "something very special will occur in the answer options for the FACTS and OPINIONS questions. You will be asked to listen carefully to the following information because it virtually ensures success on this TECAT section. Of the four answer options, two will be fact and two will be opinion almost. Of the two fact statements, one will appear in the passage and one will not appear in the passage. Of the two opinion statements, one will appear in the passage and one will not appear in the passage. This information has some very important implications...First of all, if you are asked to identify a fact, the first thing you can do is simply ignore the two opinion statements. With the two remaining, pick the one you think you have to determine which one of these two happen to appear in the passage. The same would work for the opinion statement..."

Of course the examiner still had to do the basic sorting, still had to recognize fact and opinion. But if a teacher were having trouble making the necessary distinction, wouldn't it make it easier to know that there were always two of each?

The extent of teaching to the test varied greatly. The video-taped pre- sentation had the greatest proportion of content teaching; in the worst case, content was taught less than half of the workshop day. Although content was emphasized in most of the workshops, the widespread availability of "test-taking tricks" has to be considered as a partial explanation for the extremely high passing rate. Table 1.1 shows the percentage of teachers who passed the test in each of the 45 sessions.

**Table 1.1 Performance on the TECAT: March 1994 Administration**

<table>
<thead>
<tr>
<th>Ethnic Group</th>
<th>Number Tested</th>
<th>Number Passing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hispanic</td>
<td>24,945</td>
<td>23,194 (92.3%)</td>
</tr>
<tr>
<td>Black</td>
<td>15,481</td>
<td>12,802 (83.1%)</td>
</tr>
<tr>
<td>White/Other</td>
<td>195,505</td>
<td>154,838 (79.2%)</td>
</tr>
<tr>
<td>TOTAL</td>
<td>210,284</td>
<td>150,076 (71.5%)</td>
</tr>
</tbody>
</table>

**By Age Group**

<table>
<thead>
<tr>
<th>Age</th>
<th>Number Tested</th>
<th>Number Passing</th>
<th>Percent Passing</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-29</td>
<td>36,971</td>
<td>34,971 (94.8%)</td>
<td></td>
</tr>
<tr>
<td>30-39</td>
<td>74,706</td>
<td>70,617 (95.0%)</td>
<td></td>
</tr>
<tr>
<td>40-49</td>
<td>52,349</td>
<td>50,239 (96.1%)</td>
<td></td>
</tr>
<tr>
<td>50-59</td>
<td>29,985</td>
<td>28,311 (94.9%)</td>
<td></td>
</tr>
<tr>
<td>60-69</td>
<td>4,616</td>
<td>4,055 (87.7%)</td>
<td></td>
</tr>
<tr>
<td>Over 60</td>
<td>721</td>
<td>721 (100%)</td>
<td></td>
</tr>
<tr>
<td>Not Given</td>
<td>384</td>
<td>384 (100%)</td>
<td></td>
</tr>
</tbody>
</table>

Data reported by institutions granting college degrees showed considerable variability. For example, graduates from the University of Texas at Austin, at San Antonio, and at Austin passed the TECAT at rates exceeding 90%, whereas numerous colleges in Texas (mostly small private institutions) had passing rates as low as 55%. Graduates of out-of-state colleges had a 97.9% passing rate. Early in our study, one union spokesman suggested that it was this extreme variability in the quality of graduates from different institutions, especially traditionally black colleges, that had been an underlying impetus for the TECAT.

By the second administration of the TECAT raised the first passing rate to 99%. Of the 6,579 teachers who failed the test in March, 1,704 were an equivalent version of the test in June and passed; only 1,190 teachers failed a sec-ond time; 67% teachers did not sign up to be retested. We heard vague talk about teachers who decided to retire early rather than face the pressure of the test; but the number of retirees was not appreciably higher in the districts surveyed.

---

 educational researcher
In an effort to understand what kinds of teachers had been weeded out by the test, we relied first on accounts from personnel directors and the representative sample of teachers. We were seeking corroborating evidence that the deficiencies of the failed individuals had been known in the school or district. Three points support the argument that the test got rid of incompetent teachers: nearly all teachers indicated that the TECAT was a fair test of literacy skills essential for good communication in the classroom; teachers admitted that the one or two failures they knew personally used poor grammar in their day-to-day conversation, and personnel directors classified the teachers who had been fired after failing twice, as "average" teachers—they were neither exemplary nor very bad. Only rarely (accounting for less than 5% of the failed teachers known to the respondents), did we hear of a woefully incompetent teacher who should have been fired years ago; "the test finally got him."

On the negative side, some teachers were fired because of the test who should not have been. We examined data for districts where the passing rate was 85% or less on the first testing. These districts fell into one of three categories: districts comprised of group homes for the mentally retarded or emotionally disturbed, heavily im-pacted minority districts (3 districts); or small rural districts with fewer than 30 teachers altogether. Teachers of mentally retarded children were also identified by several personnel directors as those they had lost because of TECAT. As a group, vocational education teachers had a disproportionately high failure rate. Several personnel directors noted that the loss of these individuals seemed particularly unfair because, to be certified, they had never been required to be college graduates. Often teachers who were interviewed expressed regret over a shop teacher who had been fired, "I know he doesn't speak proper English, but he really knows his machinery, and he's so good with the kids." Similar antiviva-lence regarding the legitimacy of the test versus the value of a college degree was expressed about many physical education teachers and coaches about bilingual education teachers and among the undergarment teachers. As part of the cost analysis discussed later, TEA supplied data which corroborated the overrepresentation of special education teachers, vocational education teachers, and coaches among the failures.

Teacher Morale and Public Confidence
Before our first visit to Texas, we had read newspaper accounts about teacher protests and about how insulated teachers felt at having to take a literacy test. Early on, leaders of teachers' organizations told us extensive stories about the anxiety and disruption the test had caused. But we had also been told by TEA staff and legislative aids that these stories were mostly unconfirmed. Legislators who had sponsored the reform legislation believed that militant union leaders did not speak for the majority of teachers. They knew teachers personally back home who were quite willing to take the test, if it were to get rid of the few incompetents who were giving the profession a bad name.

Interviews with scientific samples of teachers and personnel directors were intended to give a more representative picture. But even without the filters of politics or media sensationalism, we were told consistently that the test had created tremendous stress and bitterness. Most compellingly, the 25-26% of teachers who did not themselves feel threatened by the test nonetheless described its negative impact on the majority of their colleagues, stimulating the majority of personnel directors said that the TECAT had had no real effect in their district because virtually everyone passed, but it had generated negative attitudes and made teachers feel degraded. "We had a very bad year."

Teacher interviews resulted in more than 1,000 pages of transcripts: more than 100 pages were in response to these questions.

Did preparing for the TECAT make you a better teacher? Did preparing for the TECAT have any negative effects on your teaching during the past year? (end later in the interview) What has been the effect of the testing program on teachers? Only 5% of educators said that preparing for the TECAT had made them better teachers or administrators. Half said that studying or worrying about the TECAT had hurt their teaching because of the time it took, because of the stress they were under, or because they were less willing to give time to extracurricular activities.

The following verbatim quotations typify the reported effect of the testing program:

- I think mostly what I found negative was the way I saw some of my peers and also some of my superiors, those who I looked up to and respected, older people, become con-cerned and a little bit worried about the thought of having to take a test to enable you to secure your position.
- ...some of my fellow teachers just went straight into fight about the test, they were so anxious about it. I think for no reason, in some cases. There it had a negative effect on their teaching, and, as a result, I watched their frustration build, and I began to wonder what the test was about.
- The morale really dropped. I have never heard so many teachers say, "If I could find another job," "It's about time for me to retire." Things like that. It's really hurtful.
- I don't know how else to say it except it was just a humiliating experience. One of the worst experiences I may have ever been through as a teacher... I begrudge the time that I had to take to study for it, to worry about it.
- We usually get really enthusiastic about differences until we're teaching, things we're going to study and things like that, and we just couldn't get ourselves up for the test, we felt we were incompetent, and the kids, even in first grade, they would comment.
- Everybody felt like [we] were incompetent— from the kids to the governor. It was deflating to us. And everybody just felt really, really down. I think the governor's idea was that he was going to prove us competent, so teachers would go for him. Well, it backfired on him.
- I think it has given a kind of a bitter at-titude. And, I think it's been kind of just an embarrassment. You know, they've had sample questions on the test, and they're so simplistic that it's almost a joke.
- Low morale. It just really tasted them in the stomach. Many people were very nervous, uptight, concerned about it, especially the ones who had taught for over 20 years. "What will happen if all of a sudden I'm found unsuited, unfit, without credibility?"

The above responses represent over 2000 words of the teachers interviewed. Others said there was no effect on teachers or gave a positive reply such as the following: "I don't know, but I think that there definitely were some people, I include myself, who were pressed to learn too quickly, and I think that's a good thing. There were a lot of bad things, especially for those who pro-
bably know just about everything they had to know about the language. I'm pretty sure that everybody went through a very stressful time.

To gain some perspective on the negative feelings reviewed, we should note that nearly 70% of teachers said that the reading and writing skills measured by the TECAT were prerequisite to being a good teacher. "If you can do these things, you shouldn't be in the classroom." Thus, there was a large discrepancy between feelings about the underlying principle, that all teachers should be literate, and the test itself. But, to be required to take a reading and writing test when they already held college degrees made teachers feel less, not more, professional. To make matters worse, students and letters to the editor permitted in calling the TECAT a competency test, when even the governor admitted it was only a literacy test. Humiliation and embarrassment occurred because media publicity invariably portrayed teacher protests alongside examples that made the test seem laughably easy. High anxiety was created because so much was at stake; many, many teachers said that feelings would have been different if failing the test meant taking a college refresher course rather than losing their job. The pervasiveness of these themes in the two representative samples led us to conclude that TECAT consumed the attention of educators in Texas for the 1985-1986 school year and that it had a devastating effect on teacher morale.

Perhaps a more serious effect of the TECAT was its harm for public opinion. Many teachers felt that the test and accompanying publicity actually worsened public confidence in education. In fact, teachers were about equally divided on this issue. Approximately half the teachers interviewed said that the test had done what legislators had intended, it had gotten rid of the bad teachers and proven that the majority are competent. The other 50% of teachers felt that all teachers had been made to seem less competent and that the 99% pass rate made the whole thing a joke.

The following sets of quotations characterize the two conflicting positions:

Teacher Position One: The TECAT proved to the public that teachers are competent:

"Not that's where I think something good has happened. I think the public realized maybe a little bit more about how hard teachers work and really that they're bright people. I think that it could show the public that we, the majority of teachers in Texas are not illiterate. Well, it may have been a set-up in our country. Teachers probably take the brunt of society's ill. And the fact that 99% did well. And I saw some people really are always complaining that their kid's teacher's probably the stupidest person on face of the earth. I think that the public was surprised to find so many teachers could pass the test. Perhaps they think a little higher of teachers.

In the school district that I'm teaching in, most of the parents and the public had confidence in us to begin with. I think it's important that we let the public know that we are good teachers and that we are teaching then something. If the children are dropping out or they're not learning anything, it's not our fault, it's their fault.

Teacher Position Two: The TECAT was a joke. Norke the public has a lesser opinion of teachers than before:

"It was ridiculous. They think this was all a farce. I think they were hoping that this was going to weed out incompetents. Every student, including myself, has had a bad teacher. And I think they were hoping in one broad sweep they could eliminate those who were not as professional as they should be. And I think they were disillusioned because of the publicity afterwards—the way the press chose to characterize the type of test it was and how easy the test was. I think it's negative. The teachers were behind the 8-ball. If they didn't do well on the test, then Obviously they weren't good teachers. And if they did well on the test then the legislature—the first thing they said was, 'Wow! This test was too easy; we passed too many.' And, therefore, you were dead if you did and dead if you didn't. The public doesn't like the classroom teacher. People feel that we are nationally, that it was wrong for us to be angry over a test. They feel we are always whining, wanting more money. Some of the jokes circulating was consisted of several pages of ridiculous, very ridiculous things. Say, for instance, 'find your way through a maze and the way is outlined very dark.' It was trying to emphasize how dumb the questions were, but I don't think they were that dumb... You know, it was funny for a joke but it's not funny for teachers, it really demeans the teachers.

I don't think it's had any effect whatsoever on the general public. I mean, half of them are saying, 'Gee, that was ridiculous. It was a waste of money. I knew they'd do bad.' The other half are saying, 'Gee, that was a waste of money. The test was too easy so I knew they'd all pass.'

Trying to assess what the public really thinks is difficult. Even with representative survey data the results are at times internally inconsistent. In 1985, 54% of Texas gave their schools a grade of A, B, or C, a much higher percentage than in the Gallup Poll nationally. In the same survey, however, 79% said they were in favor of competency testing for teachers (The Texas Poll, 1986). The tone of most newspaper articles was consistent with the more pessimistic account. In the commentary, we would discuss how "both sides" contributed to negative publicity that may have harmed the public esteem of teachers.

Cost Analysis

How much did it cost to test every teacher and administrators in Texas? The contracted cost was $4,833,000 in special funds to develop, administer, and score the TECAT. In addition, the Texas Education Agency subsidized the appropriation by assigning regular assessment staff to the project for an estimated cost of $223,500.

A summary of the public costs of TECAST is shown in Table 2. The largest cost was the use of a teacher's time to take the test. We have also approximated the publicly sponsored costs of providing preparation for the test.

The total tax supported cost of TECAST was $33.5 million. The analysis summarized in Table 2 should be considered conservative in that costs were only included if there were data to support the estimate. Additional costs were likely occurred without our being able to represent them in the analysis.

Realistic cost data are important, however crudely estimated, because actual costs were an order of magnitude greater than the anticipated cost of testing. TECAST was expected to cost about $25 million. This was the number found in the cost estimates of the Select Committee and was still the
figure used when the State Board dis-
cussed the feasibility of implementing the legislation by testing every teacher with a Texas-developed test. In most cases Select Committee staff and Com-
troller's staff who computed cost estimates for proposed reforms included increases for both state and local jurisdictions. This was not done, how-
ever, in the case of teacher testing, therefore costs such as closing school on TECAT Monday were not included. A one-time test for practicing teachers was considered to be one of the cheapest of all the likely reforms. Data on the real public cost of TECAT indi-
cated that it was an expense more on the order of a programmatic intervention such as a proposed pre-kindergarten for disadvantaged four-year-olds, rather than an inexpensive infilling within the error of the estimates for major reforms.

In our original report we also included estimates of more than $42 million in private costs, including "wages" for teacher study time (at an average of 12 hours), and workshops, materials, and score reports purchased by teachers. The private cost analysis was con-
troversial, not because the figures were inaccurate but because it was con-
sidered inappropriate to add personal expenses (especially study time, which had no required an outlay of (cash) to an analysis of public policy decisions. In evaluation research, however, it is desirable to assign dollar values to

| TABLE 2 | Summary of TECAT Public Cost Analysis |
| Test development and administration: | |
| Nominal cost | $4,833,000 |
| TEA, staff | 232,500 |
| Teachers' inservice day to take test (20,000 teachers at $100/day) | 2,000,000 |
| Local school cost in supplying test sites (108 test days at $125/site) | 13,500 |
| Preparation workshops and review: Costs to districts and Education Services Centers | |
| Inservice development or district paid for workshops | 3,000,000 |
| Teachers' inservice x $25 | 75,000 |
| Information services and staff time (210,000 teachers x $5) | 1,050,000 |
| Site fee (300 site days x $25/site) | 7,500 |
| TOTAL PUBLIC COST | $35,614,000 |

The 887 fringes represent success in the intent to remove incompetent teachers from the classroom. The average salary paid to all certificated personnel in Texas in 1985-86 was $23,765. If this amount is increased by 20% to allow for benefits, then it could be said that the annual cost of these 887 incompetent teachers is over $25 million. As a result of TECAT, this amount of tax-
payer dollars will no longer be wasted, and hence is a savings which compen-
sates for the public expenditure on the test. Furthermore, it can be argued that the fringes represent recurring savings (less so as retirements and normal at-
trition would occur), but TECAT was a one-time expense.

Commentary

The positive case. The affirmative side for recertiﬁcation of incompetent teachers has been presented at a meeting of the American Education Research Associ-
ation by test author, Dr. James Popham, and by Texas Commissioner of Educa-
tion, William Kirby. Popham (1987) argued that society has a right to certi-
tify the competence of professionals, especially when children, as clients, have no option but to accept the ser-
vice of the education professional assigned to them. The stakes are high.

"A younger assigned to a poor teacher as a first grader may become a poor reader at second grade, and perhaps a life-long low achiever. . . . Every child in America's public schools has the right to be taught by a literate teacher." Teachers who cannot read and write adequately themselves serve as terri-
ble role models. They are not likely to require written assignments from stu-
dents. They cannot comprehend stu-
dent work, nor can they upgrade their own knowledge by reading professional literature. "Being able to read and write is clearly not the same thing as being able to teach, however, it is in our view a necessary property." Commissioner Kirby (1987) described the need for educational reform. Texas was at the "bottom of the barrel in the ration at risk." Six hundred million dollars in increased teacher salaries were fed to the test which would assure the public that teachers in the classroom could at least read and write. This proportion of the state's 16,000 teachers no longer in Texas classrooms who were there before (ap-
proximately 2,000) who failed twice or
failed once and did the take, and 8,000 who never showed up to take the test). Kirby asserted that the public at-
titude toward education has improved and legislatures are more willing to
spend money on education, even in the face of a $2 billion deficit.

The negative case. The case against
backlog was so great that by
1988, 90% of all elementary
school districts had at least one
teacher in each grade level.

The problem was not the lack of
teachers, but the lack of
competent teachers. Kirby pointed out
that the test taken by teachers
was not the same as the test taken by
students. The test taken by teachers was
administered by the state, while the test
 taken by students was administered by
private companies.

The conclusion was that the test was
biased and unfair, and that it was
biased against minority students.

A few years later, a similar
test was given to teachers in the
same district. The results were
similar, but the teachers who failed
the test were not punished, and
the district did not try to
improve its test-taking
performance. Kirby argued that
the district was lazy and
unwilling to improve its
performance. The district
was also criticized for
attitudes towards minority
students.

The problem was not the lack
of teachers, but the lack of
competent teachers.

The solution was not to
increase the number of
teachers, but to improve
the quality of the teachers.

The test validity, passing, and failing.
The argument for a literacy test for
recertification of current teachers is
really a dispute about validity. Is there
a level of reading and writing profi-
cency that is so fundamental that
the test that is used to measure it
is adequate? The test is supposed to
measure reading and writing profi-
cency, but it is not clear that
the test is adequate to measure
this.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.

The test is a multiple-choice test.
It is not clear that this is an ade-
quate way to measure reading and
writing proficiency.
Iron worry. The humiliation arose because a literacy test, not a profession-
ally uplifting exam, was the source of the anxiety. To be imprisoned in study-
ing for a basic skills test that would determine someone's had a demoraliz-
ing effect, even though the perceived threat was only a matter of keeping with the in-
tention of the test.

Negative portrayal of teachers:

Pampered by the disagreement among educators about the effect of TECAT on public education, a systemic content analysis was conducted of two complete newspaper files to examine the portrayal of teachers.

Teacher incompetence in Texas had been a recurring education story culmi-
nating in early 1981 legislature time entering and exiting teacher candi-
dates. It was a dominant theme dur-
ing Select Committee hearings. Perot made headlines with the charge that.

"The Dumbest People in College Study to Be Teachers" (Januarey Globe,
5/17/84).

A second theme in newspaper stories had to do with the unprofessionalism of teachers. They were portrayed as self interested and anti reform. Perot called education lobbyists "pick-
pockets" who went taxpayers to "send more money to a zip short" (Ameri-
can Statesman, 4/4/84). Later in the special session, teacher groups were cast as mercenaries when they lobbied against equalization. By protesting the test on the basis of their lifetime con-
tractual rights, teachers' unions ap-
peared to defend incompetence. After having agreed to the reform package, three of the teacher groups withdrew in a huff, which legislators denounced as capricious and unethical.

In the year leading to TECAT, teachers were in the news filing law-
suits, studying for the test, picketing the governor, and complaining about acute report labels and proctor qualifi-
cations. Interviewed after the test, indi-
vidual teachers told reporters that it was insidiously easy. In stories that followed, legislators complained that the test had been too easy and the passing rate too high.

We cannot assert what individuals citizens believed about the TECAT. We can say that for a period of three years, news stories conveyed an undistinguished and anti-education image to teachers.

Informally, one of the intentions of the reformers had been to raise the

steem of the teaching profession. The Select Committee had been dismayed by a survey of 1983 Texas honor gradu-
ates; only 12% said they would con-
sider teaching in public education. The status of teaching was to be raised by increasing salaries and by removing in-
centives, let alone because of the posi-
tive tone of the test and the behavior of teacher organizations, young stu-
dents might now see teaching as less admirable profession than before.

Cost benefits. The public cost of TECAT, counting taxpayer supported review sessions, was over $55 million. The real cost was roughly $45 million.

original estimates, which pictured teacher testing as a cheap reform.

The direct effect of TECAT was to eliminate approximately 2,000 of Texas' 210,000 teachers. TECAT also had the effect of drawing approximately 180,000 teachers into review of rules of grammar. By all accounts, however, 90% of teachers were not in need of such review.

Unanticipated side effects of the testing program were the negative impact on moral morale and, potentially, the harm done to the public image of teaching by negative publicity.

The public and private costs of TECAT represent opportunity costs that should be weighed against other ways of achieving the same ends or expend-
ing the same resources. For example, how might the investment in profes-
sional development have been spent, if not on TECAT? What other policies might have fostered the removal of 1% of the teaching force judged to be in-
competent? What if the state had set up a fund to support the administra-
tive and legal costs of dismissing bad teachers? (For an account of the nor-
mal procedures and impediments to removing incompetents, one of which a feared legal costs, see Bridges, 1986).

Postscript. When the study began in the spring of 1986, legislation to test practicing teachers was pending in two states and talked about in others. To-
day, these actions are neither passed nor pending. We sense, as does Dr. Chris Phipps who monitors these things, that the Education Commission of the States, that there is less enthusiasm to jump into teacher testing now than 18 months ago. We can speculate that the waning interest is due to economic change, to the passing of the first wave

of reform, or to foreknowledge about the intensity of teacher union reac-
tions. Perhaps an informal network of legislators has purveyed the story that TECAT cost Texas Governor Mark

White, his job. In any case, our re-
search suggests that there is sense to legislators' caution. Although the blame for negative side effects must be widely shared, the realities of teacher testing deny the simplicity of the in-
tended policy: "to give a test and elimi-
nate the few teachers with indubitably weak communication skills.

'Way agrees with the Contentions that an admi-
tional 10,000 teachers who never showed up to take the test can be claimed as additional suc- cesses for the TECAT. Many in this number reflect normal attri-
tion. Based on teacher interview data, some very able teachers left the profession because they were felted by the presence of TECAT. This latter con-
tention is consistent with national data showing that teachers leaving the profession are on average higher scoring and better educated than those remaining.


American Psychological Association, American Educational Research Association, National Con-

A study of the Texas teacher. Technical Report. Los Angeles: Center for Stan-
dard setting, Evaluation, and Standards, Univer-
dard setting, Evaluation, and Standards, Univer-
dard setting, Evaluation, and Standards, Univer-
dard setting, Evaluation, and Standards, Univer-
dard setting, Evaluation, and Standards, Univer-
dard setting, Evaluation, and Standards, Univer-
dard setting, Evaluation, and Standards, Univer-