

California State Auditor

B U R E A U O F S T A T E A U D I T S

STAR Program:

Ongoing Conflicts Between the State Board of Education and the Superintendent of Public Instruction as Well as Continued Errors Impede the Program's Success



April 2000
99131

The first copy of each California State Auditor report is free.
Additional copies are \$3 each. You can obtain reports by contacting
the Bureau of State Audits at the following address:

**California State Auditor
Bureau of State Audits
555 Capitol Mall, Suite 300
Sacramento, California 95814
(916) 445-0255 or TDD (916) 445-0255 x 216**

OR

This report may also be available
on the World Wide Web
<http://www.bsa.ca.gov/bsa/>

Permission is granted to reproduce reports.



CALIFORNIA STATE AUDITOR

MARY P. NOBLE
ACTING STATE AUDITOR

STEVEN M. HENDRICKSON
CHIEF DEPUTY STATE AUDITOR

April 5, 2000

99131

The Governor of California
President pro Tempore of the Senate
Speaker of the Assembly
State Capitol
Sacramento, California 95814

Dear Governor and Legislative Leaders:

As requested by the Joint Legislative Audit Committee, the Bureau of State Audits presents its audit report concerning our evaluation of the State Board of Education (board), the Superintendent of Public Instruction (superintendent), and the California Department of Education's (department) implementation and administration of the Standardized Testing and Reporting (STAR) program. The State uses this program to assess the academic achievement of public school students in grades 2 through 11.

This report concludes that ongoing conflicts between the board and the superintendent about their respective roles and responsibilities have negatively affected all aspects of the STAR program. Implementing the program has been especially difficult because, due to these conflicts, the superintendent has never developed an annual implementation plan, as state law requires. Moreover, despite having a contract for the spring 2000 test cycle that better clarifies the responsibilities of the board and department, the test publisher continues to receive conflicting instructions.

The test publisher's performance in the first two years of the STAR program—spring 1998 and 1999 test cycles—has been problematic. Missed deadlines, unreliable data, and inaccurate reporting of achievement test results have plagued the program. Unrealistic timelines that the Legislature imposed for designating an achievement test did not allow the board time to thoroughly investigate potential publishers before making a final selection. Also, the superintendent did not fulfill a legal requirement to establish a formal method for working with the test publisher and as a result there appears to have been very little monitoring of the test publisher's performance. These factors contributed to the test publisher's unsatisfactory performance.

Some action has been taken by the Legislature, board, department, and test publisher to improve the implementation and administration of the STAR program. But further improvements must be made to ensure the success of the Public Schools Accountability Act, which places greater emphasis on the accuracy, completeness, and timeliness of the STAR achievement test results. It requires these results to be included in the State's Academic Performance Index (API) that the department will use for the first time in fall of 2000 to determine how to distribute funds earmarked for schools and teachers. Currently, the results of the achievement test are the sole component of the API and, therefore, the only determinant of these monetary awards.

Respectfully submitted,

MARY P. NOBLE
Acting State Auditor

BUREAU OF STATE AUDITS

555 Capitol Mall, Suite 300, Sacramento, California 95814 Telephone: (916) 445-0255 Fax: (916) 327-0019

CONTENTS

<i>Summary</i>	1
<i>Introduction</i>	5
<i>Audit Results</i>	
Ongoing Disputes Between the State Board of Education and the Superintendent of Public Instruction Undermine the STAR Program	11
Recommendations	25
<i>Responses to the Audit</i>	
State Board of Education	29
<i>California State Auditor's Comments on the Response From the State Board of Education</i>	33
California Department of Education	35
<i>California State Auditor's Comments on the Response From the California Department of Education</i>	43

SUMMARY

Audit Highlights . . .

Our review of the State Board of Education's (board), superintendent of public instruction's (superintendent), and the California Department of Education's (department) implementation of the Standardized Testing and Reporting (STAR) program disclosed:

- Open conflict between the board and the superintendent as well as errors on the part of school districts and the test publisher have negatively affected the program.*
- The superintendent has not developed an annual implementation plan, as law requires.*
- Time pressures imposed by the Legislature made it difficult for the board to thoroughly evaluate potential test publishers and select a final test.*
- During the first two test cycles—spring 1998 and spring 1999—the department did not closely monitor the performance of the test publisher. The program has been plagued with missed deadlines, unreliable data, and inaccurate reporting of achievement test results.*
- The department must take further action to ensure the success of the Public Schools Accountability Act of 1999, such as pushing for better test security.*

RESULTS IN BRIEF

The implementation of the Standardized Testing and Reporting (STAR) program in California's public schools has been problematic due to open conflict between the State Board of Education (board), which is appointed by the governor, and the superintendent of public instruction (superintendent), an elected official. The decades-old conflict over each party's respective roles and responsibilities for administering the California Department of Education (department), as well as continued errors on the part of the school districts and the test publisher has negatively affected the development, administration, and implementation of the STAR program.

Implementing the program has been especially difficult because, due to this conflict, the superintendent has not developed an annual implementation plan. State law requires the superintendent to submit to the Legislature a plan for producing valid, reliable, and comparable individual student scores. Ideally, this plan would clearly outline the board's and department's responsibilities for implementing the program, establish criteria for measuring the reliability of test results, and provide an estimate of program costs.

One example of disputes between these parties is a breakdown in communication that caused the test publisher to temporarily halt its production of testing materials and incur significant expense to resume printing the materials. Specifically, the test publisher was to field-test 16 writing questions for grades four and seven. Although the board's executive director approved all the questions, the department approved only 14. The test publisher began production of the field-testing materials with 14 questions for each grade until the board pointed out that the contract required 16 questions for the field test. To resolve this conflict, the publisher halted production and then submitted additional questions to the department. The department approved two additional questions for each grade, and the publisher resumed production in late January 2000, at significant added expense.

The Legislature's wish to adopt a single statewide achievement test as quickly as possible has created additional problems for the STAR program. To speed the process for selecting a standardized achievement test, the Legislature imposed time pressures on both the superintendent and the board to evaluate potential test publishers and select a final test. State law enacted in October 1997 gave the superintendent three weeks to evaluate candidates and send forward a recommendation to the board. The same state law gave the board a little over a month to have its independent experts evaluate candidates and designate an achievement test. This unrealistic timeline made it difficult for the board to thoroughly evaluate the quality and capabilities of the publishers under consideration; it also did not allow the board to thoroughly evaluate the final candidate's procedures for monitoring its subcontractors, contingency plans, and computer capacity.

The compressed deadlines for selection of the test and preparation of the STAR program have had a cost: The test publisher's performance during the first two test cycles of the program has been deficient in certain areas. The department did not closely monitor the publisher's performance during these years. Missed deadlines, unreliable data, and inaccurate reporting of achievement test results plagued the program. The department contends that it did not have the legal authority or resources to monitor the publisher's performance because, as required by law, the test publisher had contracted directly with the school districts, not with the department. However, state law specifically requires the superintendent to establish a method of working with the test publisher, so the superintendent did have the authority to monitor the publisher's performance. Nevertheless, the superintendent did not develop a clear description of the work to be done by the publisher, establish a schedule of completion for major activities and milestones, create a process for monitoring the publisher's performance, or define roles and responsibilities for the department, board, and test publisher.

Some action has been taken by the Legislature, board, department, and test publisher to improve the management of the STAR program since the first two test cycles. To improve the department's oversight of the publisher, the department now contracts directly with the publisher. The test publisher has also sought guidance from its consultant on improving its operations. But further improvements must be made to ensure the success of the Public Schools Accountability Act of 1999, which creates a greater responsibility for the department, school districts, and test publisher to ensure the accuracy, completeness, and

timeliness of the achievement test results. This act requires the superintendent to include the test results in the Academic Performance Index (API), a mechanism that will determine how the State will distribute about \$150 million to schools and teachers. The API will be implemented for the first time in fall 2000.

Currently, the achievement test results are the sole component of the API and, therefore, the only determinant of these monetary awards. To ensure the integrity of the testing process and maximum fairness for the students assessed, the school districts, the publisher, and the department must push for better test security. One way to do this is to require all school districts and testing personnel to attend the publisher's training sessions to be sure that they properly follow the step-by-step procedures for administering the achievement test and securing test materials.

RECOMMENDATIONS

To facilitate communication between the board, superintendent, and the department and create a more productive environment for the STAR program, the Legislature should establish a mechanism for appointing a mediator to resolve disputes that will most certainly continue concerning these entities' respective roles and responsibilities.

To ensure that the STAR program is successful, the board, superintendent, and department must fulfill their designated responsibilities and improve their policies and procedures in the following ways:

- With the assistance of a mediator, the board and the department should establish a memorandum of understanding that outlines their respective roles and responsibilities for implementing the STAR program.
- The superintendent should develop an annual implementation plan for producing valid, reliable, and comparable individual student scores as mandated by law. Further, the plan should explain how to communicate instructions to the test publisher.
- The department must continue its weekly meetings with the test publisher, as outlined in the current contract. It should also ensure that it places similar requirements in all future

contracts. The board and department must establish a formal meeting schedule to make sure that the board is kept abreast of ongoing program issues.

- The board and department should review the recommendations of the publisher's consultant and amend the current contract to ensure that the publisher does implement all recommendations that will improve the STAR program.
- Finally, the department should calculate the additional costs of requiring all school district and testing personnel to attend training courses on properly administering the test and accurately reporting necessary demographic information. If the costs are reasonable in relation to the total program costs, the department should take the necessary actions for requiring all relevant personnel to attend this training.

AGENCY COMMENTS

While the board agrees with our findings, it generally does not agree with our recommendations. Specifically, the board believes that the governance issue between it and the department is much broader than the STAR program and to the extent that a solution is found, it should be established in this broader domain. Furthermore, while the board agrees that it needs to ensure that the publisher implements those recommendations that will improve the STAR program, it does not necessarily agree that the contract must be amended to accomplish this result. However, the board does agree that a formal meeting schedule to facilitate communication between it and the department is necessary.

The department agrees with our recommendations. In addition, the department provides information concerning its view of the STAR program's implementation. ■

INTRODUCTION

BACKGROUND

The State uses the Standardized Testing and Reporting (STAR) program as just one assessment of the academic achievement of public school students. This particular testing program is used in more than 1,100 districts statewide. California first began statewide testing in 1962 using various achievement tests. Later, the State moved to a single standardized test that it administered at several grade levels to gather information at the district and state levels. In 1972, legislation created the California Assessment Program (CAP) to measure the average achievement at a particular district or school. Students at four grade levels participated in the CAP tests. However, to reduce costs and test-taking time, each student only took a portion of the test. Also, the test was not designed to produce reliable individual student scores.

CAP was used until spring 1990. A year later it was replaced with the California Learning Assessment System (CLAS), administered to students at three grade levels. CLAS differed from the multiple-choice CAP test in that it focused on open-ended questions designed to test a student's ability to complete a specific task and provide a reliable, valid score for each student. However, CLAS encountered two problems in spring 1994: Parents in some communities objected to the personal nature of some questions in the reading and writing portions, and concerns were raised about the consistency of scoring of the writing tests.

These issues led the governor, in the fiscal year 1995-96 budget, to veto a bill continuing CLAS, citing its unreliable scores for individual students and its lack of emphasis on basic skills. In October 1995, the governor signed into law the voluntary Pupil Testing Incentive Program. Under this program, districts received \$5 for each student in grades 2 through 10 who took a basic skills achievement test approved by the State Board of Education (board). Unfortunately, the variety of tests the school districts administered did not allow for easy comparison of test scores.

In October 1997, the governor signed the STAR program into law and discontinued the Pupil Testing Incentive Program. To allow for statewide comparison of scores for individual students,

schools, and school districts, each school district must test all students in grades 2 through 11 using the Stanford 9 achievement test, designated by the board. In April 1999, the Legislature passed the Public Schools Accountability Act. This act attached the achievement test scores to monetary and nonmonetary incentives for teachers and schools by creating the Academic Performance Index (API). The State has set aside up to \$150 million for schools and teachers that show performance growth in their API score from one year to the next. Currently, the achievement test scores comprise 100 percent of the API. In the future, about 60 percent of the API will be based on these and other test scores, with the balance incorporating school attendance and graduation rates, as soon as the superintendent of public instruction (superintendent) deems these indicators reliable.

ROLES AND RESPONSIBILITIES FOR THE STAR PROGRAM

The board is appointed by the governor with the advice and consent of two-thirds of the Senate and is composed of 11 members. The board's staff consists of an executive director, deputy executive director, staff counsel, program analyst, education program consultant, and support staff. Responsibilities of the board include adopting policies to implement the program, approving the contract between the California Department of Education (department) and the test publisher, and evaluating the performance of the publisher at the end of each test cycle. In addition, legislation enacted in 1995 required the board to adopt statewide content standards designed to establish the academic knowledge, skills, and abilities that all public schools are expected to teach—and students are expected to learn—for reading, writing, math, history/social science, and science. The board finalized its content standards for these subject areas in October 1998. The legislation also required the board to adopt performance standards to measure whether a student, school, or school district has met the content standards. The board is required to complete its adoption of performance standards by mid-November 2001. The superintendent, who is elected by the people, is responsible for executing the policies the board establishes.

Responsibilities for administering the STAR program are shared among the department, school districts, and test publisher. The department is responsible for day-to-day administration. The school districts must ensure that the test is administered in accordance with state laws and regulations as well as with the

test publisher's instructions. The test publisher is responsible for developing the test, compiling test scores, and producing reports for individual students, schools, and districts for statewide comparisons by the statutory deadline. The Figure on the following page outlines each entity's administrative responsibilities under the program.

SCOPE AND METHODOLOGY

The Joint Legislative Audit Committee (JLAC) asked the Bureau of State Audits to conduct an audit regarding the implementation and execution of the STAR program. The JLAC requested us to analyze the contracting procedures of the board and the superintendent and to assess the performance of the board, department, school districts, and test publisher in implementing, administering, and reporting the results of the STAR program.

To determine whether the board properly handled the selection process for the test publisher, we reviewed the applicable laws, regulations, and pertinent documents, such as the request for publishers to submit a test, the publishers' proposals, underlying documentation of the superintendent's recommendation of a publisher submitted to the board, and the evaluations of the board's independent experts, as well as applicable correspondence and board minutes. We also interviewed the staff of the board and the department.

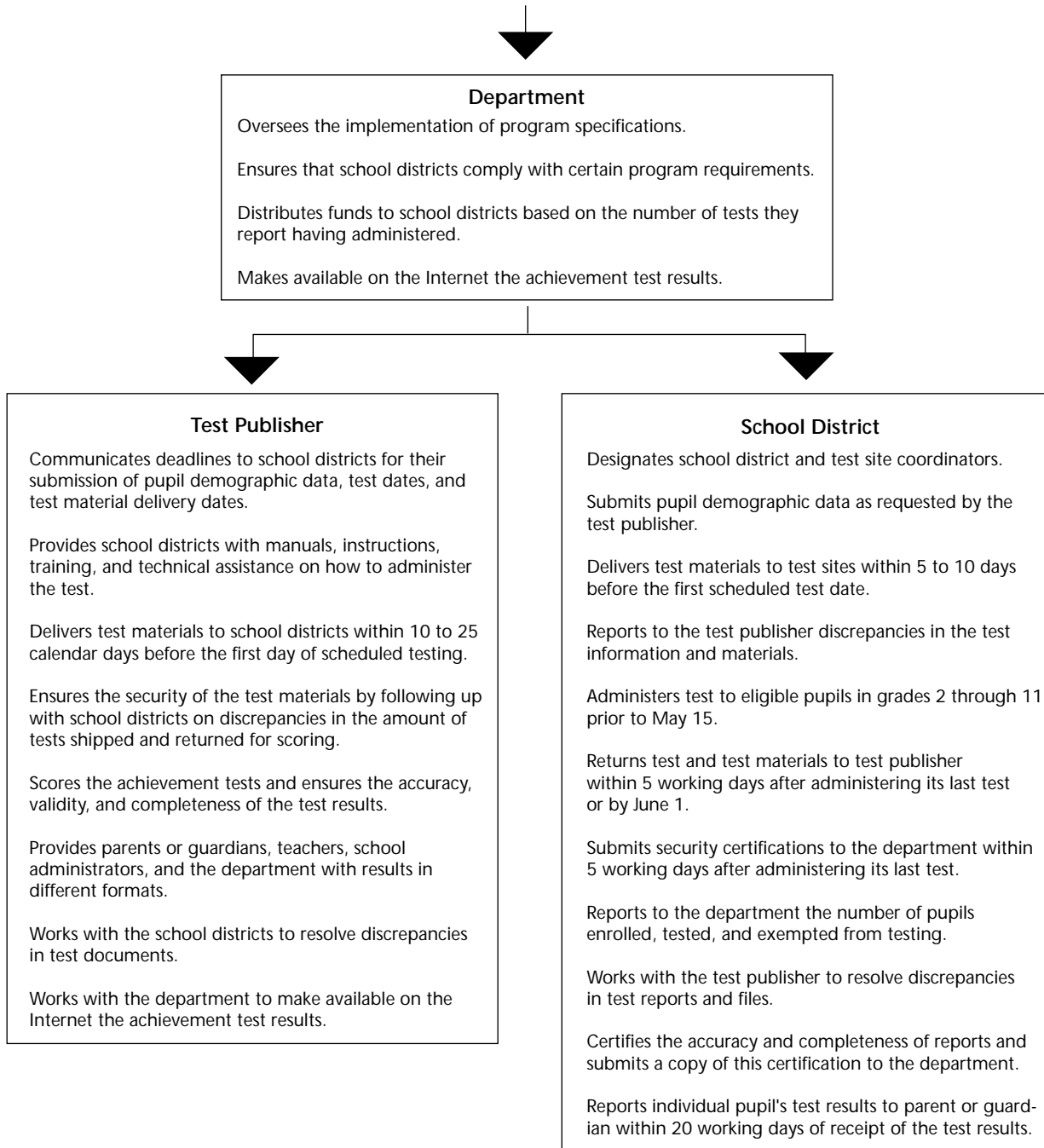
We ascertained the roles and responsibilities of the board, department, school districts, and test publisher for implementing the STAR program. We also assessed whether administrative policies and procedures for handling the program were adequate and whether important deadlines were met. We determined whether clear and effective communication existed between the board, department, and test publisher to address program issues. Moreover, we assessed the process the department and school districts followed to execute the contracts with the test publisher in the spring 1998 and spring 1999 test cycles. Finally, we examined the training that school districts and their testing personnel received.

To determine the causes of—and the responsibility for—the late delivery of test materials and of the reporting problems in the first two test cycles, we reviewed relevant policies, procedures, and agreements. We also surveyed 54 school districts, as well as reviewed documents such as board minutes, publisher progress

FIGURE

Three Entities Play a Role in Administering the STAR Program

The administration of the **STAR testing program** requires the coordinating efforts of the California Department of Education (department), the test publisher, and school districts. The primary goal of this program is to assess students' ability to solve problems and communicate their skills in subject areas such as reading, spelling, written expression, and math.



reports submitted to the board, and correspondence between the board, department, and test publisher. We then reviewed a November 1999 report from a consultant evaluating the publisher's operational procedures, controls, and quality assurance policies. The test publisher requested this review to identify breakdowns in its operation that delayed delivery of test materials and caused errors in its reporting of test results during the spring 1999 test cycle.

To determine what corrective actions the board, department, and test publisher have taken and to evaluate whether these actions are sufficient to prevent future problems, we examined the current year contract and determined the progress the test publisher has made in implementing its consultant's recommendations. We also attended state board meetings; conducted interviews of the staff of the board, department, and test publisher; and examined documentation to determine whether these parties took appropriate and timely actions and if any further difficulties had surfaced in the spring 2000 test cycle. Finally, we determined the potential impact that late and inaccurate STAR achievement test results could have on the API. ■

Blank page inserted for reproduction purposes only.

AUDIT RESULTS

Ongoing Disputes Between the State Board of Education and the Superintendent of Public Instruction Undermine the STAR Program

SUMMARY

Historically, the State Board of Education (board) and the superintendent of public instruction (superintendent) have had ongoing disputes concerning their respective roles and responsibilities for administering the California Department of Education (department). Because they do not agree on their duties in implementing the Standardized Testing and Reporting (STAR) program, the board and the department often issue conflicting instructions to the test publisher. No single mechanism is in place that allows the board and department to resolve their disputes concerning what constitutes an administrative versus policy issue for the STAR program. Until a mechanism is put in place that allows these parties to clarify and agree on their respective responsibilities, their disputes will continue to undermine the program.

AMBIGUOUS ROLES OF THE BOARD AND THE SUPERINTENDENT CREATE CONFLICT BETWEEN THE BOARD AND THE DEPARTMENT

The California Education Code (code) gives the board the authority to adopt policies for the governance of kindergarten through grade 12 in public schools. The code further states that the role of the superintendent and the department is to administer the board's policies. The board and the superintendent, however, have not always agreed whether certain issues are matters of policy or administration.

In 1849, California's first Constitution established the elected position of the superintendent of public instruction, and in 1852, the Legislature established the first State Board of Education. The governor, the state surveyor, and the superintendent

sat on the board. Its role was to advise and supervise the superintendent on apportioning school funds to the counties. Subsequent constitutional amendments gave the governor authority to appoint the board and expanded the board's policy-making role. The board operated largely in an advisory capacity until 1912, when the California Constitution was amended to require the Legislature to provide for the appointment or election of the board. This constitutional amendment, and the statutes the Legislature enacted to implement it, gave two bodies authority over education in the State.

As early as 1919, the department correctly predicted that giving two distinct bodies responsibility for education could cause conflicts.

As early as 1919, the department correctly predicted that giving two distinct bodies responsibility for education could cause conflicts. In 1926, open conflicts between the board and the superintendent did indeed surface. Unable to reconcile their differences over governing the actions of the department, the board and the superintendent have, over the years, turned to the attorney general to clarify their roles. In 1943, the attorney general concluded that the department was jointly administered by a governing and policy-making body, the board, and its executive officer, the superintendent. He compared their relationship to that of the executive head of a corporation and its board of directors. He also suggested that the Legislature clarify ambiguous language related to the powers, duties, and functions of the board, the superintendent, and the department.

The Little Hoover Commission (commission) has also commented on the conflicts between the board and the superintendent. In a 1990 report, the commission concluded that the superintendent had assumed the role of policy maker rather than operating at the direction of the board. The commission indicated that ambiguities in state law and the California Constitution concerning the roles and responsibilities of both parties contributed to this problem. The commission also found that the governance structure for education is inherently flawed because it requires a policy-making board appointed by the governor to have an elected official carry out its policies. The commission also commented that because the superintendent is accountable to the people instead of the board, the individual elected to that position can choose whether to follow the board's policies, thus reducing the power of the board's direction to mere suggestions.

As the conflicts continued, the Third Appellate District of the Court of Appeals of California [*State Board of Education v. Honig*, 13 Cal. App. 4th 720 (1993)] considered whether the board or

the superintendent was responsible for setting education policy. The court held that “the Legislature clearly envisioned a department administered jointly by the board and the superintendent.” The court clarified that the board is the policy maker. The board’s role is to establish goals affecting public education and methods for achieving them, as well as principles to guide the department’s operations. While the board is to “direct and control” the department, it is not to “micromanage” it. The superintendent’s role is to supervise department staff in the day-to-day execution of the board’s policies and oversee the more detailed aspects of the programs and the budget.

Following the court’s decision, legislators have introduced various bills to further address the governance issues facing the superintendent and the board, but none have passed. Senate Bill 856 of the 1991-92 Regular Session attempted to overrule the court’s decision and to vest primary policy-making authority with the superintendent’s position. The governor, however, vetoed the bill, reasoning that governors can have influence over education policy only if the board maintains control over policy decisions. More recently, a bill analysis prepared by the Assembly Education Committee on Senate Bill 839 of the 1999-2000 Regular Session, which sought to clarify the board’s and superintendent’s roles, suggests that the joint legislative committee created in May 1999 to establish and develop a “Master Plan for Education–Kindergarten through University” would likely review the issue of state governance of K–12 education, thus further postponing any clarification of the board’s and superintendent’s roles.

THE STAR PROGRAM LACKS AN IMPLEMENTATION PLAN

The Honig case gave the board and the superintendent some guidance about their respective roles in state governance of education, but the decades-old conflict between these educational bodies continues and has negatively affected all aspects of the STAR program. Implementing the program has been especially difficult because, due to this conflict, the superintendent has not developed an annual implementation plan.

State law requires the superintendent to submit to the Legislature a plan for producing valid, reliable, and comparable individual student scores. Ideally, this plan would clearly outline the board’s and department’s responsibilities for implementing the program, establish criteria for measuring the reliability of test

The decades-old conflict between the board and superintendent has negatively affected all aspects of the STAR program.

results, and provide an estimate of program costs. Department staff contend that the board would prevent them from implementing any plan the superintendent developed.

Despite ongoing disagreements between the board and the superintendent concerning their respective roles and responsibilities, the superintendent has a legal requirement to prepare the implementation plan. Although such a plan might not prevent the board and superintendent from challenging each other's respective roles and responsibilities, it could provide a blueprint for this program and address issues such as defining the criteria for measuring the reliability of test results and estimating costs.

POOR COMMUNICATION CONTINUES DESPITE A CONTRACT THAT CLARIFIES SOME RESPONSIBILITIES

The department contracted directly with the publisher, with the approval of the board, for the spring 2000 test cycle. This contract expands the department's responsibilities, indicates which tasks require board approval, and provides more detail on the publisher's methods and timelines for accomplishing its tasks. Despite this contract, weak communication between the department, the board, and the publisher continues. To remedy miscommunications between the department and the publisher, the contract requires the department to hold weekly conferences with the publisher to discuss ongoing issues. The department and the board also began holding weekly meetings in mid-January 2000 to allow the two parties to discuss current issues before communicating with the test publisher.

Although allowing the publisher to contract directly with the department instead of the school districts was a significant change, the board and department continued their disputes and gave the publisher conflicting instructions several times during the spring 2000 test cycle. In one instance, a breakdown in communication between the board and department caused the publisher to temporarily halt its production of field-testing materials and incur significant expense to resume printing the materials.

In this particular case, the publisher had developed 16 questions for use in a field test for grades four and seven and, following a review by two separate panels, submitted the questions for final approval to both the department and the board. The board's

Several times during the spring 2000 test cycle, the board and department gave the publisher conflicting instructions.

executive director made some modifications but approved all of the questions. In contrast, the department requested further information regarding the development of the questions, modified several of the questions, and then deleted two questions per grade. The test publisher began production of the field-testing materials with 14 questions for each grade until the board insisted that the contract required 16 questions for the field test. The publisher halted production and then submitted additional questions to the department. The department approved two additional questions for each grade, and the publisher resumed production in late January, at significant added expense.

When they discovered the discrepancy in their actions, the board's executive director and the director of the department's Standards and Assessment Division (division) examined their roles in the approval of test questions. The board's staff questioned the department's decision to wait until the final approval to modify and remove items rather than present its concerns earlier when the questions were before the review panels. The division director questioned whether it was appropriate for the board to approve the questions, viewing this approval as an administrative function that should not involve the board. To complicate matters, it is unclear why the board was involved in the approval process, because the contract requires approval from the department, not the board.

Conflicts and poor communication are also evident in the test publisher's development of a supplemental math test for the spring 2000 test cycle. The supplemental math test includes newly developed questions as well as questions from the existing basic math test. In 1999, the supplemental math test that the publisher developed to conform to California's content standards produced results that did not meet the industry standards for reliability. The board and the department discussed how to raise the test to the desired reliability level. Although the board, department, and test publisher disagree on the details of these discussions, it is apparent that at one point, a board member suggested that the publisher add five questions to the supplemental test to boost its reliability. However, due to the difficulty of finding more questions that met the State's standards, the task of adding five math supplemental items was not included in the contract.

In 1999, the math supplemental test produced results that did not meet industry standards for reliability.

In December 1999, the publisher decided to pursue the addition of the five questions. The board's staff determined that the five additional questions should be reviewed by the panel of testing

experts and presented at the board's January 13, 2000, meeting. The test publisher and board staff also discussed the issue with the department. Nevertheless, during the board meeting, department representatives and several board members strenuously objected to adding the five questions. The department stated that it had not had an opportunity to assess the proposed questions, and some board members stated that they had not reviewed any documentation enabling them to make an informed decision. The board did not take any formal action on the issue.

Because additional questions were not added, the board asserts that the current math test is not reliable enough to compare students' or school districts' performance.

The department later requested more information on the questions from the publisher, which the publisher provided. The publisher warned the department that it needed a decision on the questions by January 28, 2000. After this date, any changes would pose a serious threat to the programming and quality assurance procedures for the STAR program. However, collaboration between the test publisher and the department broke down. The department charged that the test publisher did not provide all requested information in time, while the publisher claimed that the department did not ask for the information in enough time for the department to complete its analysis. Because the test publisher decided it could not add the questions, the time and energy spent on this issue were wasted. More importantly, board staff told us that without the extra questions, the present supplemental math test is not reliable enough to use to compare students' or school districts' performance. The test publisher states that using results from the supplemental math test would be problematic because most educational measurement professionals become increasingly uncomfortable with grouping scores from tests that do not meet the industry standards for reliability. In its February 2000 meeting, the board directed the publisher to develop a plan to ensure the necessary reliability.

These examples raise several questions concerning the parties' roles and communication. First, it is questionable whether the board should have considered the questions before the department had a chance to judge their technical merits. While representatives of the board and the publisher stated that the department had ample notice of potential board action on this issue, the department believes that it was not allowed to fully participate in the decision-making process. It is also apparent that communication between the test publisher and the department was not adequate to allow them to agree on the additional questions by the deadline of January 28, 2000. The above situations underscore the need for the board and the department to resolve disputes

over administrative or policy issues before communicating with the test publisher. Both board and department staff acknowledge that they do not agree on the level of board involvement necessary to implement the program. These conflicts will continue, and the test publisher will continue to find itself between two state entities issuing conflicting instructions until a third party is designated to mediate disputes between the board and department.

SHORT TIME FRAMES AND THE SUPERINTENDENT'S FAILURE TO ADEQUATELY EVALUATE THE TEST PUBLISHER'S PERFORMANCE CONTRIBUTED TO PROBLEMS WITH THE PROGRAM

Unrealistic timelines that the Legislature imposed on the board for designating an achievement test and the superintendent's failure to establish a formal method for working with the test publisher that included structured monitoring contributed to the test publisher's unsatisfactory performance during the spring 1998 and spring 1999 test cycles. The Legislature's desire to adopt a single statewide achievement test placed great time pressures on the superintendent and the board and did not allow the board time to thoroughly investigate potential publishers before making a final selection. Furthermore, the superintendent's failure to monitor the publisher chosen allowed problems to continue during the first two years of the program.

In August 1997, the superintendent urged the governor and the Legislature to consider the drawbacks of adopting an achievement test too quickly. Despite this concern, legislation to establish the STAR program was approved by the governor on October 9, 1997, and took effect immediately. The legislation required school districts to administer an achievement test, which the board would select, to each student in grades 2 through 11 beginning in spring 1998. It also required the superintendent to recommend an achievement test to the board by October 31, 1997, and required the board to select a final test no later than November 14, 1997.

The Legislature imposed timelines that made it difficult for the board to thoroughly evaluate each interested party's credentials before designating its final selection.

Despite the superintendent's warning, the Legislature mandated a very short time frame for the selection process, making it difficult for the board to thoroughly evaluate each interested party's credentials before designating its final selection. This compressed timeline gave the superintendent only about three weeks to invite test publishers, testing companies, and other interested parties to submit achievement tests; assemble a

representative group of education professionals to examine and compare the tests' content and technical criteria; and evaluate the findings and forward a recommendation to the board. Similarly, the board had slightly more than a month to obtain and review information on test publishers, organize a review panel of independent experts to evaluate test publishers' materials and review their comments, and designate an achievement test.

State law governed the board's method of evaluating each test publisher. The board assembled a panel of independent experts

and requested test publishers to submit materials to the independent experts, as well as to board members, by October 27, 1997. The independent experts reviewed the tests using the criteria set forth by state law and commented on the strengths and weaknesses of each publisher's ability to meet the legal criteria (see text box). The superintendent used a similar process to evaluate information obtained from the test publishers.

In addition to the information obtained from the superintendent and its independent experts, the board arranged for each publisher to present its materials and answer questions during a November 1997 board meeting. The board based its final decision mainly on the publisher's cost estimate and the publisher's abilities to produce a test that met the State's curriculum requirements and to supplement the test with additional questions—once the board was able to adopt content standards for certain subject areas. Content standards are the specific academic knowledge, skills, and abilities that public schools are expected to teach—and students at each grade level tested are expected to learn—for reading, writing, math, history/social science, and science.

Although content standards are critical to the testing process, other critical factors also warrant a thorough review. Two of the board's independent experts made clear the need for the board to consider additional information prior to selecting the current test publisher. One reviewer advised

the board to determine if the publisher's estimated costs would cover the services and technical support that the STAR program would require. Another independent expert cautioned the board

Criteria for Selecting a Test Publisher

- Can the publisher produce valid, reliable individual student scores?
- What is the quality and age of empirical data supporting the national norm used in the test?
- Can the publisher meet the State's requirement for reporting results by June 30?
- Can the publisher report results that can be compared with the results of previous standardized achievement tests administered by the school districts?
- Is the publisher able to provide results comparable with test results from the 1998 benchmark year and subsequent years?
- Is the publisher able to align the test with the content and performance standards adopted by the State Board of Education?
- Has the publisher forfeited a performance bond?
- What is the estimated per-student cost of administering the test?
- What are the test publisher's procedures for ensuring the security and integrity of test questions and materials?
- Does the test publisher have experience in successfully conducting testing programs in other states?

about the publisher's use of numerous subcontractors, which would require more coordination to be sure the work was running smoothly. However, the board did not consider these or other issues, such as the capability of the selected test publisher's computers to handle the anticipated test volume, its contingency plans for handling mishaps, and its procedures for monitoring its subcontractors' performance.

Although required to do so by law, the superintendent did not establish a method for working with the test publisher to ensure that test results are valid, reliable, and comparable.

Equally important, there appears to have been very little monitoring of the test publisher's performance by the department in the first two test cycles, spring 1998 and spring 1999. State law requires the superintendent to establish a method for working with the test publisher to ensure that the achievement test results are valid, reliable, and comparable, but the superintendent did not establish a formal method for working with the test publisher in the first two years. Thus, a clear description of the scope of the work; a timeline for major activities and milestones; a plan for monitoring the publisher's performance; and defined roles and responsibilities for the department, board, and test publisher did not exist. Although state law dictated timelines for some activities, such as designating the test, establishing the testing period, and reporting test results, the law did not establish timelines for other critical functions, such as the development and approval of program specifications for scoring tests.

The department contends that it did not have the resources or legal authority to monitor the publisher's performance because the test publisher's contractual relationship was with the school districts. However, state law specifically requires the superintendent to establish a method of working with the test publisher to ensure that valid, reliable, and comparable test results were produced, and thus the superintendent had the authority to monitor the publisher's performance.

Perhaps if the board had more time to consider the additional factors and the superintendent had established a formal method of working with the test publisher, the implementation of the STAR program would have been smoother. The test publisher's performance in the first two years was problematic, particularly during the spring 1999 test cycle. As a result, in August 1999, the board reduced the amount the publisher was paid by \$1.1 million, and state law now allows the department to impose a penalty of up to 10 percent of the cost for any component the publisher fails to adequately deliver.

In its attempt to improve its performance, in July 1999, the test publisher obtained the services of a consultant to identify breakdowns in its operations and those of its subcontractors. These breakdowns had delayed delivery of test materials and caused errors in the reporting of test results during the spring 1999 test cycle. Based on the consultant's report released in November 1999, we question the publisher's preparedness to handle the STAR program. The consultant found that delays of up to two months in computer modifications for the publisher and a subcontractor held up delivery of test materials to some school districts. According to the test publisher, its computer system required tailoring to address changes in the spring 1999 test cycle and its subcontractor's system required customization to meet program needs. Delays in making these changes did not give the publisher and its subcontractor sufficient time to test their systems and identify design flaws before they actually processed orders from the school districts.

These delays prevented the subcontractor from receiving orders earlier and meeting its planned schedule for shipping test materials to school districts at a rate of 100,000 tests per day. Instead, it shipped test materials at a considerably lower rate of 65,000 tests per day. To meet the shortfall, the subcontractor added another computer as well as more equipment and personnel. When we asked what measures the test publisher had taken to ensure the subcontractor could meet the required volume, the test publisher stated that it had relied on its

subcontractor's verbal assurance and its past experience with the subcontractor that the subcontractor's equipment had the necessary capacity. Moreover, when we asked the test publisher if a contingency plan existed to address potential mishaps resulting from computer or shipping problems, the test publisher stated such a plan did not exist.

In response to complaints from school districts that did not receive test materials within 10 days prior to their first day of testing and were unable to obtain information from the test publisher's staff on delivery dates, the department conducted a survey of 292 school districts in March 1999 to

determine the extent of the problem (see text box) and worked with the school districts and the test publisher to resolve the delivery delays. The school districts responded that the delays created many problems. The districts were forced

Summary of Responses From 292 School Districts to the Department's Survey

- 145 (50 percent) did not receive materials on time.
- 58 (20 percent) were not given a revised delivery date.
- 111 (38 percent) had to reschedule testing for some or all of their schools.
- 17 (6 percent) reported that they would be unable to test some of their students.

to cancel or reschedule other activities originally scheduled around the achievement test, including field trips, spring breaks, assemblies, and district proficiency and performance assessments. Further, district staff had to work overtime to prepare the testing materials.

The test publisher assures the board and department that it is attempting to correct its previous shortcomings. Its consultant made several recommendations, such as improving channels of communication between the test publisher and its subcontractors and creating additional quality control procedures. The test publisher is implementing its consultant's recommendations, but there is one recommendation for which more work must be done. Specifically, the consultant recommended that the test publisher automate its quality control procedures for verifying the information in its Internet report files, but the publisher is still investigating ways to automate its quality control process more fully. The consultant was scheduled to perform a follow-up review in February 2000 to evaluate the progress made by the test publisher; however, the consultant did not receive the publisher's approval to proceed. The publisher states that this follow-up review will proceed after the spring 2000 test cycle.

TEST RESULTS ARE CRUCIAL TO ADDITIONAL SCHOOL FUNDING

The Public Schools Accountability Act of 1999 places great emphasis on the accuracy, completeness, and timeliness of the STAR achievement test results. It requires these results to be included in the State's Academic Performance Index (API), which the department will use for the first time in fall 2000 to determine how to distribute about \$150 million earmarked for schools and teachers. Currently, the results of the achievement test comprise 100 percent of the API. Using the test results as the sole criteria for distributing funds creates a high-pressure, "high-stakes" environment. Teachers and schools are placed under greater public scrutiny and responsibility for their students' performance. Therefore, it is imperative that school districts, schools, and teachers ensure the integrity of the testing process and the accuracy of the information they provide to the test publisher. Further, the department must ensure that it produces the achievement test results by July 15, as state law now requires, so it can publish the API and distribute the funds on schedule.

Using the test results as the sole criteria for distributing about \$150 million to schools creates a high-pressure, "high-stakes" environment for both teachers and students.

Errors Plague Program Data and Delay Mandated Reports

Since the inception of the program, school district and publisher errors have prevented the department from posting complete and accurate test results on the Internet by June 30.

For the first two years of the STAR program, school district and publisher errors prevented the department from posting complete and accurate test results for public viewing on the Internet by the yearly statutory deadline of June 30. The test publisher is required to produce statewide reports as well as reports for each grade, school, district, and county. The publisher must also produce reports grouping the scores by the students' language proficiency, gender, and economic status.

Although the publisher submitted the first year's reports by the deadline, these reports were missing demographic data for more than 750,000 (18 percent) of public school students. To be able to compare schools statewide, the department requires school districts to provide demographic information such as the student's date of birth, grade, gender, ethnicity, and language proficiency level to the test publisher. School districts have the option to purchase the test publisher's pre-identification (pre-ID) service, which allows them to receive student answer sheets with this information and saves time for students and teachers on the day of the test. Mistakes made by one of the subcontractors contributed to errors in the 1998 reports. Because the test publisher's subcontractor did not know the proper procedure for correcting pre-ID documents, the scanning feature on its computer system simply overrode the pre-ID information with the manual changes the school districts supplied.

In addition, the 1998 reports that the test publisher prepared for the school districts contained 698 instances of misreported information affecting about 8 percent of California's 8,500 schools. Misreporting occurred because the codes that some school districts entered on their data forms to identify special instructional programs were different from the codes that the department assigned. The test publisher did not have adequate quality control procedures to ensure that the districts used only assigned codes on the data forms. Consequently, scores for some schools were not reported, reported under the wrong school codes, or reported for schools that do not exist. It took the publisher almost six months to work with the school districts and resolve these errors to the satisfaction of the board.

Although the test publisher took measures to avoid repeating these errors during the second year of the STAR program, new problems cropped up in the June 30, 1999, reports. In these reports, the test publisher erroneously categorized the scores of

certain students proficient in English with students of limited English proficiency. According to the test publisher, this mistake occurred because of unclear reporting specifications and its inability to perform adequate quality control in this area. Additionally, for the second year of the testing program, the publisher's staff misinterpreted the coding system and reported the norming codes for 44 year-round school districts as though they were traditional school districts. Norming codes specify the averages that the test publisher should apply while scoring the tests and are based on the number of instructional days for each school district. Because the number of instructional days can vary at a year-round district, an incorrect norming code can produce inaccurate data. This error, which affected the scores for about 190,000 students, was not detected until the school districts reviewed their reports. Consequently, the test publisher had to revise its reports for the department and school districts.

Delays in posting accurate and complete results can have a significant effect on the API.

The test publisher was able to correct the reports in both years, but the 1999 reporting errors delayed the department's mandated posting of results on the Internet by at least 22 days. Delays in posting accurate and complete test results can have a significant effect on the API. The department plans to publish the 2000 API data in September and award payments to schools and teachers in October or November. Due to the time it takes to calculate the API, the director of the department's Office of Policy and Evaluation believes that any delay in reporting test results may delay the calculation and reporting of the API. More importantly, if the department posts incorrect test results, schools and teachers may not receive the monetary awards they deserve. This will in turn affect students, as the department will grant schools that demonstrate performance growth in their API score greater latitude in spending any new or existing school funds to further improve their students' performance.

In addition to its financial effect on schools, the API can have unintended effects throughout the State. Parents may perceive the API—which is now solely based on the achievement test results—as an indicator of the effectiveness of their children's schools and make decisions that could shift enrollment in certain school districts and affect school funding. In light of these possible and serious consequences, the department must ensure that future test results are complete, accurate, and available by July 15 as state law now dictates.

Properly Training School Districts to Administer the Tests Can Increase the Integrity of the STAR Program and the Accuracy of Its Data

The use of the STAR achievement test results to determine the monetary awards for schools and teachers creates a “high-stakes” assessment environment that holds individuals and schools responsible for their performance. High-stakes assessment and the relationship between test scores and money available for education generally increase public scrutiny of test results, the pressure on students and teachers or schools to perform well, and the need for better test security to ensure maximum fairness for students taking the tests. Consequently, the State must ensure the integrity of the testing process by making sure all personnel involved in the STAR program receive adequate training to properly administer the tests.

The API creates a greater need for the State to ensure the integrity of the testing process.

School districts and test site coordinators are responsible for following the test publisher’s procedures for securing the test materials, monitoring their use, and returning unused materials. However, the test publisher stated that not all of the school districts returned their unused test materials, which compromises the integrity of the testing process. School districts and test site coordinators are also responsible for ensuring that the demographic information submitted to the test publisher is accurate, complete, and timely. Districts that opted for the pre-ID service during the spring 1999 test cycle were required to submit their demographic data for each student tested to the publisher between 15 and 30 days before the scheduled delivery date for the test.

Of the 246 school districts opting for the pre-ID service and scheduled to begin testing in March 1999, 67 (27 percent) submitted pre-ID files that were either late or contained errors. Also, during the spring 1998 test cycle, some school districts that purchased the test publisher’s pre-ID service allowed students to use test materials designated for others or provided conflicting demographic information on their pre-ID forms. These types of errors affect the accuracy of the data used to group and compare student and school test results. To encourage schools to submit accurate information, the board amended its regulations to enable the department to withhold \$1.25 of the \$2.52 per-student payment to districts for each incomplete or incorrect student answer sheet submitted to the publisher. However, if the school districts, schools, and teachers are properly trained on the importance of the information and how to accurately report it to the test publisher, the likelihood of errors will decrease.

Each year, the test publisher makes available to school district and site coordinators training that includes discussions on the proper procedures for securing test materials and reporting demographic information. For the spring 1999 test cycle, the test publisher held training at 26 locations and distributed materials that clearly outlined all procedures that school district and site coordinators must follow to administer the achievement test. Further, the test publisher emphasized that school districts must ensure that site coordinators receive this training and, in turn, that site coordinators ensure that teachers, examiners, and proctors also receive this training.

Training school districts, schools, and teachers on the importance of accurately reporting information to the test publisher decreases the likelihood of errors.

Despite the strong urging for the site coordinators to go to the publisher's training sessions, as well as personnel who assist them, there is no requirement for them to do so. Of the 46 school districts responding to our telephone survey, 12 (27 percent) stated that they do not require their teachers, test examiners, or proctors to receive training. Moreover, 6 of these school districts stated that their district coordinators did not attend the publisher's training sessions. Of the remaining 34 school districts that did train their staff, 23 stated that there was no additional cost because they added it to their existing training, 2 stated that there was a minimal cost associated with sending their district coordinators to the sessions, and 3 incurred costs ranging from \$300 to \$44,000. The training costs will vary for school districts depending on their size, the number of staff attending, and travel expenses.

With the new element of high stakes placed on test results, the department and the school districts must increase their emphasis on ensuring the integrity and accuracy of the results. Thus, the department should ensure that all school staff responsible for administering the achievement test attend the test publisher's training sessions. Although a mandatory training requirement may increase the program's cost, the benefit derived from this training will improve the credibility of the testing process and the reporting of the achievement test results.

RECOMMENDATIONS

To facilitate communication between the board, superintendent, and the department and create a more productive environment for the STAR program, the Legislature should establish a

mechanism for appointing a mediator to resolve disputes that will most certainly continue concerning these entities' respective roles and responsibilities.

To ensure that the STAR program is successful, the board, superintendent, and department must fulfill their designated responsibilities and improve their policies and procedures in the following ways:

- With the assistance of a mediator, the board and the department should establish a memorandum of understanding that outlines their respective roles and responsibilities for implementing the STAR program.
- The superintendent should develop an annual implementation plan for producing valid, reliable, and comparable individual student scores as mandated by law. Further, the plan should explain how to communicate instructions to the test publisher. The plan should include:
 - A decision matrix that indicates the representatives that must be present from each entity before a decision is accepted.
 - Timelines indicating all anticipated actions to be taken by the board and the department.
- The department must continue its weekly meetings with the test publisher, as outlined in the current contract. It should also ensure that it places similar requirements in all future contracts. The board and the department must establish a formal meeting schedule to make sure that the board is kept abreast of ongoing program issues.
- The board and department should review the recommendations of the publisher's consultant and amend the current contract to ensure that the publisher does implement all recommendations that will improve the STAR program.
- Finally, the department should calculate the additional costs of requiring all school districts and testing personnel to attend training courses on properly administering the test and accurately reporting necessary demographic information. If the costs are reasonable in relation to the total program costs, the department should take the necessary actions for requiring all relevant personnel to attend this training.

We conducted this review under the authority vested in the California State Auditor by Section 8543 et seq. of the California Government Code and according to generally accepted government auditing standards. We limited our review to those areas specified in the audit scope section of this report.

Respectfully submitted,

A handwritten signature in black ink that reads "Mary P Noble". The signature is written in a cursive style with a large, prominent "M" and "N".

MARY P. NOBLE
Acting State Auditor

Date: April 5, 2000

Staff: Joanne Quarles, CPA, Audit Principal
Nasir Ahmadi, CPA
Juan R. Perez
Joel Riphagen

Blank page inserted for reproduction purposes only.

Agency's comments provided as text only.

John B. Mockler
Executive Director
California State Board of Education
721 Capitol Mall
Sacramento, CA 94244-2720

March 24, 2000

Ms. Mary P. Noble
Acting State Auditor
California State Auditor
555 Capitol Mall, Suite 300
Sacramento, CA 95814

RE: State Board of Education and the California Department of Education: Ongoing
Conflicts Between These Entities and Continued Error Impede the STAR
Program

Dear Ms. Noble:

This letter is the State Board of Education's response to the above-referenced audit report recently conducted by your office at the request of the Joint Legislative Audit Committee. The response consists of a brief background of events leading up to the current STAR program and a response to each recommendation that is directed to the State Board of Education. The State Board has not responded to those recommendations that are directed to the California Department of Education. Additionally, the State Board has provided additional factual information relevant to the audit report's discussion of the development of writing prompts.

Background

The State Board of Education has been engaged in carrying out a standard and assessment reform of K-12 education in California. This effort began with the adoption of academic content standards in 1997 and 1998 and the implementation of a statewide testing program in 1998. The change to a standards-based educational system with accountability for both students and schools has not been without opposition. Academic issues that were debated during the adoption of the academic content standards have continued to be debated in the development and implementation of the Standardized Testing and Reporting (STAR) Program. The enactment of the STAR program was not without controversy and although the STAR test is in the third year of administration, there are still critics of the use of standardized testing in California. The continuing debate over standards and standardized testing has made it difficult to implement the Board's policy decisions in the STAR program. This difficulty has proven to be most troublesome in the area of mathematics.

Response to Recommendations

The audit report recommends that:

“The Legislature should establish a mechanism for appointing a mediator to resolve disputes that will most certainly continue concerning these entities’ respective roles and responsibilities.”

The report further recommends that:

“With the assistance of a mediator, the board and the department should establish a memorandum of understanding that outlines their respective roles and responsibilities for implementing the STAR program.”

Although the audit report is thoughtfully prepared and accurate in most instances, the State Board does not agree with the recommendation that a mediator be appointed to resolve disputes between the State Board and the State Superintendent with respect to appropriate policy and administrative roles. The Legislature and the Governor have the authority to establish the system of governance for K-12 education. The current system vests policy decisions with the State Board of Education, whose members are appointed by the Governor. The administration of the system carried out through the California Department of Education (CDE) under the oversight and direction of the State Superintendent of Public Instruction.

The debate over the appropriate governance structure to govern education has raged for almost a hundred years and recently has been the subject of several reports including recommendations by the Constitutional Revision Commission. There are bills pending in the Legislature to revise the governance structure and the Legislature has had hearings in recent months on this topic. The issues involved in this discussion are much broader than the STAR Program. To the extent that a solution is found, it should be established in this broader domain.

①*

The audit report further recommends that:

“The board and department should review the recommendations of the publisher’s consultant and amend the current contract to ensure that the publisher follows through with implementing those recommendations that will improve the STAR program.”

In January, the State Board directed the publisher to address the issues raised by the publisher’s consultant. In the context of that discussion, the publisher assured the Board that the consultant’s recommendations were being implemented. The State Board agrees with the audit report recommendation that the State Board follow-up with the publisher to ensure that the publisher implements the necessary recommendations from the

*California State Auditor’s comments appear on page 33.

consultant's report. However, at this point, the State Board does not necessarily agree that the contract must be amended to accomplish this result.

The final recommendation applicable to the State Board, directs that:

"The board and department must establish a formal meeting schedule to ensure that the board is kept abreast of ongoing program issues."

In January the State Board's executive director and CDE's chief deputy responsible for the Standards and Assessment Division implemented a weekly testing meeting to facilitate communication between the State Board and CDE regarding relevant testing issues not only in the context of the STAR Program but all statewide testing programs.

Development of Writing Prompts

There is also one portion of the report with which the State Board takes exception. It may be helpful to have a more detailed explanation of events. Specifically, the report describes an event in which 16 writing prompts were delivered to both the State Board's executive director and the director of the CDE Standards and Assessment Division. The State Board's executive director approved the writing prompts with some modification whereas the CDE director of the Standards and Assessment rejected two prompts. The State Board acknowledges the audit finding that there were miscommunications between the State Board staff and CDE staff with respect to the approval of the writing prompts. Moreover, if an administrative mechanism for the acceptance or rejection of contract deliverables had been in place perhaps these events described in the audit report could have been resolved more efficiently. However, at the same time, the State Board does not believe that the dual approval requested by the publisher caused the publisher's production delay.

The publisher had been told repeatedly both verbally and in writing that the approval of contract deliverables was not final until the CDE made the approval. This understanding was placed as part of the contract approved by the State Board. More importantly, the publisher was under a contractual obligation to develop not 16 writing prompts but 24 prompts to allow for the extensive community review process. However, although the publisher did develop the required number of prompts, the publisher only delivered 16 of the 24 prompts to CDE. As soon as the State Board staff learned that CDE staff had rejected two prompts and that the publisher had decided to field test only the 14 approved writing prompts, the State Board staff immediately notified CDE that the reduction to 14 prompts was an amendment of the contract that had not been approved by the State Board. CDE staff agreed with the State Board staff and subsequently required the publisher to produce the additional writing prompts with the end result that 16 prompts currently are being field-tested. As these events unfolded the publisher did halt production as stated in the audit report.

②

The State Board does not dispute that its approval of the content of the writing prompts was not required. The production delays arose because the publisher did not provide the full number of prompts required by the contract. The required number of prompts is a contract issue, which under the terms of the contract is subject to the approval by the State Board's executive director. Unfortunately there was no administrative process in place for the timely acceptance or rejection of contract deliverables that could have alerted CDE staff to these discrepancies and facilitated the resolution of the discrepancies in a more timely manner.

The auditors assigned to review this program worked diligently and should be commended for their professionalism and diligence in what must have been a challenging assignment. If the State Board can be of any further assistance, please contact John Mockler, the State Board's Executive Director at (916) 657-5478.

Sincerely,

(Signed by: John B. Mockler)

John B. Mockler
Executive Director

COMMENTS

California State Auditor’s Comments on the Response From the State Board of Education

To provide clarity and perspective, we are commenting on the response to our audit from the State Board of Education (board). The numbers below correspond to those we have placed in the board’s response.

- ① The board is correct in stating that there have been debates over—and various attempts to address—California’s governance structure for K-12 education for almost 100 years and that this issue is indeed much broader than problems with the STAR program. The establishment of a more appropriate governance structure for K-12 education is complex and may take several years to resolve. However, we believe that there is an immediate need to address the conflicts between the board and the superintendent of public instruction concerning the STAR program. Thus far, these conflicts have negatively affected all aspects of this testing program. Therefore, we stand by our recommendation that a mediator be appointed to resolve disputes that will most certainly continue concerning these entities’ respective roles and responsibilities.
- ② While the board correctly states that the contract requires the test publisher to develop 24 writing questions, it fails to mention that the contract does not contain language that explicitly requires the test publisher to deliver all 24. Nevertheless, our report focuses only on the miscommunication concerning the test publisher’s contractual obligation to field test 16 writing questions.

Blank page inserted for reproduction purposes only.

Agency's comments provided as text only.

Delaine Eastin
State Superintendent of Public Instruction
California Department of Education
721 Capitol Mall
Sacramento, CA 95814

March 27, 2000

Mary P. Noble
Acting State Auditor
555 Capitol Mall, Suite 300
Sacramento, CA 95814

Audit No. 99131

Dear Ms. Noble:

Thank you for the opportunity to review and comment on your draft report entitled "State Board of Education and California Department of Education: Ongoing Conflicts Between These Entities and Continued Errors Impede the STAR Program." Overall, we would first like to convey that we believe the draft report's emphasis on the historical issues of governance between the past State Board of Education (State Board) and the California Department of Education (CDE) does little to convey accurately the challenges attendant with the implementation of the Standardized Testing and Reporting (STAR) program. We also find a disconnect between the summary to the draft report and the report itself. Therefore, some initial points are in order to provide a full context for the period of time which we understood was the focus of the audit: the 1998 and 1999 administrations.

①*
⑦

First, in the fall of 1997 the State Superintendent of Public Instruction (Superintendent) was directed by the Legislature and Governor to evaluate the tests submitted for possible use in the STAR program. She was given three weeks to do so. In that time, she redirected staff, convened a panel of state and national testing experts, and forwarded to the State Board a comprehensive, exhaustive analysis of the three tests submitted, along with her recommendation. This was a significant investment of time and resources from CDE. Not only did the State Board, as then constituted, pay little attention to the analysis and recommendation of the Superintendent, but the State Board also convened another, different panel of experts to do essentially the same work the Superintendent was charged by statute to do.

②

This duplication of effort set a tone for the administration of the STAR program not captured in the audit report. Since that point in time, the former State Board made clear that it had a direction it wanted to pursue for the STAR program, and the professional judgements and recommendations of the Superintendent and CDE staff were frequently

*California State Auditor's comments begin on page 43.

ignored. This was evident as the State Board and its staff developed relationships with the test contractor and appointed test development panel members to work directly with the test contractor. With limited staff resources, unrelenting time pressures, and additional important policy priorities, often the Superintendent and CDE staff had to choose the level of energy and time to invest in a program over which we ultimately had no authority, especially given the 1,000 separate contracts with local school districts.

3

Indeed, the test publisher's California representatives met frequently with State Board members, without the knowledge of CDE staff. These were meetings that CDE staff not only were not invited to, we often only learned about them after the fact. State Board members insisted their direction to the test publisher was strictly on matters of policy, over which CDE had no role. The Superintendent formally asked the State Board to discontinue the practice of having one or two State Board members drive testing policy. We provided the auditors with copies of the Superintendent's correspondence.

4

Second, plans, timeframes, roadmaps, and budgets are all good tools. Additionally, we recognize the need, on a week-to-week, month-to-month, and year-to-year basis, to identify the policy questions and issues, steps, and status of the STAR system in our work with the State Board. We have taken initial steps towards doing this with an assessment notebook presented to the State Board at its March 2000 meeting. CDE has worked each year with the Department of Finance and Legislature regarding the needs of the testing system as those needs cover growing programs costs, staffing projections, and new and existing contracts. We mistakenly believed that our communications with the Department of Finance and Legislature were sufficient for program reporting purposes under Education Code section 60604. They are not. Consequently, we are working on an annual report and will present it to the Legislature for the testing programs covered by the statute in the next few weeks.

5

However, while we acknowledge the statutory requirement to produce such a report, we question the utility of the report as suggested by and envisioned in the audit report. We believe the audit report should distinguish between the utility of internal planning documents and the utility of an annual report for assessment programs that have undergone extensive change. State testing in California has changed dramatically every year since 1994. When one considers the evolution of the testing system covered by Education Code section 60604, and the requirements to be included in the annual report, it is critical to understand that the system is not static. Long-term planning, no matter how desirable, remains something of a moving target. The Governor and the Legislature are critical players and their roles have been most significant. Consider:

1994: California Learning Assessment System (CLAS). The Legislature seeks but the governor vetoes reauthorization of the testing system.

1995: California does not have a statewide test. The Legislature passes and the Governor signs Assembly Bill 265, which authorizes new state and local testing systems,

along with academic content and performance standards development. The bill also requires the development of an annual plan by the Superintendent to be submitted to the Legislature no later than March 1 of each year. The state test (matrix) is to be developed after the approval of state content and performance standards. The local system is the origin of what would, two years later, become STAR.

1996: California begins the Pupil Testing Incentive Program (PTIP), which initially has 56 approved tests administered exclusively at the local level. School districts can choose among the tests. Development of content standards in English-Language Arts and Mathematics begins. No progress is made on the state test (matrix).

1997: PTIP enters a second year. School districts continue to be able to select from among those tests on the approved list. Concerns arise about the lack of comparability among the different tests used. The Legislature and Governor negotiate the selection of a single state test to be used in place of PTIP. No change in the status of the state test (matrix). In November, the State Board selects Harcourt's Stanford Achievement Test—Form T as the state's single test under the STAR program. Rather than a single state contract, the test publisher has standard agreements (contracts) with each school district. CDE is not given a formal role. The State Board rejects CDE's suggestion that a substantial penalty of as much as 50 percent be imposed should the test publisher fail to comply with its contract. Although other states have much higher penalties, the State Board decides on a 10 percent penalty. Content standards are approved in English Language Arts and Mathematics. Development of standards in History-Social Science and Science begins. ③

1998: STAR begins its first year of administration. Development of the augmentation (standards-based) test begins in English-Language Arts and Mathematics. No change in the status of the state (matrix) test. Content standards are approved for History-Social Science and Science.

1999: STAR begins its second year of administration. Individual contracts with school districts are dropped; a single state contract with CDE is to replace it for 2000. The Public School Accountability Act introduces new purposes for state tests. The augmented standards-based test is given for the first time. History-Social Science and Science augmented tests are developed; these tests will not be administered until 2001. The Superintendent recommends a contractor to develop the state (matrix) test; the State Board declines to take action. The Legislature and Governor direct the Superintendent to develop and the State Board to approve a High School Exit Exam. The test publisher is penalized the full amount for its failure to meet its commitments. The test publisher makes major changes in its California personnel.

As you can see, developing long-term budget and planning documents in the face of such change may be statutorily required, but the utility is limited. Further changes are in store for the testing system for 2000, including additional accountability purposes, a new ⑤

writing sample in grades 4 and 7, and the development of the High School Exit Exam, which may or may not have a long term connection to the STAR program. When the testing system attains some stability, such long-term planning will be critical to its success.

Corrections to the Report

Production of Writing Prompts for STAR 2000 Field-Testing: The development and field testing of writing prompts is occurring entirely in the scope of the 1999-2000 contract, with the prompts being developed for use in STAR 2001. Discussion in the draft report, on pages 18 and 19, addressing the conflict around approving the STAR writing prompts for the spring 2000 field-test does not include all relevant information. First, the draft report does not acknowledge that the CDE's contract with the test publisher required the publisher to develop 24 writing prompts for each grade 4 and 7. The publisher only forwarded 16 prompts for each grade to CDE. CDE made multiple requests to the publisher requesting the additional eight prompts for each grade. CDE received some additional prompts without the literature on which the prompts were based. Finally, when CDE received the four prompts available from the test publisher, CDE was able to work with the test publisher to modify additional prompts for use in the spring field-testing within the timelines established by the publisher. To date, CDE has not received the full complement of 24 prompts per grade required under the contract.

6

Second, the writing prompts that CDE removed from the field-test had elements that had been found objectionable during the CLAS tests.

Third, throughout the process of modifying additional prompts for the field-test, CDE and test publisher staff worked collaboratively to ensure that timelines were met and that the appropriate spiraling of test forms could be done. The test publisher provided no indication to CDE that work would stop on field-test form production or that additional costs would be incurred. CDE has not received documentation of any excess costs incurred by the publisher. To the best of CDE's knowledge, the field-testing of writing prompts occurred as planned.

7

Finally, the statements in the summary erroneously lead the reader to believe that the test publisher was a "victim" of the CDE. The more complete account on pages 18 and 19 of the draft report identifies the role of the State Board in development of the writing prompts.

7

Augmented Math Tests: The draft report, on pages 19 and 20, recounts a controversy over the possible addition of five questions to the augmented math test, in an effort characterized in the report "to raise the test to the desired reliability level." While the report states on page 19, "While the board, department, and test publisher disagree on the details of these discussions, ...", the narrative continues to describe a version of events that the auditors have constructed. The draft report proceeds as if the version presented is

an accurate representation of the series of events surrounding augmentation of the math test. This “incident” is used to illustrate poor communication among the State Board, CDE, and the test publisher. While CDE does not dispute poor communication, it does dispute the portrayal of events surrounding the possible addition of math items. If matters of fact cannot be established (as the draft report suggests), that alone should be adequate to illustrate the problem of poor communication. If more details were needed for the illustration, each statement of “fact” should be attributed to its source so that the reader is not led to a conclusion about “facts” that are based on selective recall of the parties involved.

8

After extensive discussion over a period of months during State Board meetings, from which no formal board action was taken, a board member and State Board staff asked the test contractor to bring forward five additional items for consideration. CDE staff were not apprised of this activity, and it is our position that such policy decisions and direction must be made by the State Board in an open, public meeting. A majority of State Board members concurred, and expressed their intent to have all policy decisions on STAR made by the entire State Board during open meetings.

9

The discussion on pages 19 and 20 of the draft report also does not mention the work completed by CDE staff to facilitate the publisher’s reconfiguration of the STAR augmented math tests for spring 2000 so that the tests would be better aligned with industry standards. When CDE analyzed the effect of adding the five recommended items, the test characteristics were not significantly improved. Because the addition of the items would not have improved the test, and their addition would have increased reliance on the Stanford 9 rather than moving toward a fully standards-based test, the items were not added to the math test.

10

Responses to Specific Allegations

CDE’s Failure to Monitor the Test Publisher Performance: Starting on page 16 of the draft report, CDE is cited for not closely monitoring the test publisher’s performance during the first two years of the STAR program. The draft report asserts that because state law requires a plan, CDE is also authorized to monitor the performance of the publisher. The allegations about the lack of a plan (as required by Education Code sections 60604(a) and (b)) and the relationship between the plan (or lack thereof) and most of the problems that occurred during 1998 and 1999 merit further comment.

11

CDE disagrees with the draft audit report on matters of authority, resources, and enforcement of compliance. CDE has (related to the STAR program) some very narrowly delineated responsibilities and authority. Those areas of responsibility are not referenced in the draft report.

11

At the time the STAR program was implemented, CDE received one analyst position to implement the STAR program. Thus, establishing a plan and monitoring the test

- 12 publisher's performance with only one analyst makes it virtually impossible to succeed. CDE provided the auditors with documentation of the positions requested each year (through the budget process) for implementation of STAR. The draft report makes no reference to these efforts or to the fact that the positions were denied each year. CDE has been forced to implement STAR using staff assigned to work on other programs. We believe it would have been helpful to have checked with other states to find out the level of staff support required for the work plan imagined by the auditors.

Even if CDE had been provided adequate support for STAR, CDE still contends it did not have the authority to fulfill appropriate monitoring activities. Accordingly all through 1998 and 1999 budget hearings, CDE repeatedly brought this issue up and requested the needed authority.

Poor Communication Between CDE, State Board and Publisher: Page 17 of the draft report states that "weak communication among the department, the board, and the publisher continues." Generally, both the CDE and the test publisher perceive communications to work when there is adherence to a single point of contact both for the CDE and the publisher. Weak communications occur when individual State Board members or State Board staff work directly with the publisher rather than working through and with CDE staff responsible for the contract. CDE is, after all, the contract monitor for the 2000 contract with the publisher.

- 13 Nonetheless, the new State Board majority has gone so far as to establish a policy that individual State Board members should not engage in direct negotiations with the test publisher and the Superintendent and her staff feel that there has been a substantial improvement in communication. Likewise, the test publisher's staff changes have helped clear the air.

The information on page 25 of the draft report seems to imply that problems encountered with the spring 1999 testing program and the penalty levied against the test publisher may have been avoided. There is no mention of the CDE work that occurred during the delivery problems. The draft report states that CDE conducted a school district survey, but there is no reference to the fact that CDE made repeated requests to the test publisher for delivery status information. CDE did eventually receive daily reports and intervened to assist in prioritizing school district deliveries. CDE also worked with school district coordinators to develop contingency plans for administering the tests if materials were received late from the test publisher.

- 14

In general, as CDE staff advised your auditors, CDE is working towards better communication between the State Board, CDE staff and test publisher.

Training for School District and Test Site Coordinators: Pages 31 – 33 of the draft report address the need for school district and test site coordinators to receive training for implementing the STAR program. CDE has no means to compel or enforce school

- 15

district compliance with statute short of legal action. CDE has no authority in statute to require school district personnel to attend either CDE or publisher workshops to prepare for the testing program. Generally, administrative action within the schools is a local issue, unless CDE is given specific authority in statute.

Nevertheless, CDE has taken several actions to strongly encourage school staff to attend training sessions. CDE conducted workshops for coordinators prior to testing and encouraged testing coordinators to attend the publisher's workshops, where they would receive detailed implementation procedures. CDE staff have been available by telephone and electronic mail to assist school district coordinators with implementation of the STAR program. Additionally, the test publisher provided all school district coordinators with guidance and a videotape clearly explaining their responsibilities. The materials provided to the school district coordinators stated that the coordinator "is responsible for ensuring that all procedures and policies required by the STAR program are implemented within the school district." The publisher provided manuals to site coordinators which stated, "Formal, classroom-style training for teachers/examiners (and any proctors who have been enlisted) is strongly recommended."

Recommendations

1. CDE welcomes Legislative assistance in clarifying State Board and CDE roles and responsibilities.
2. CDE agrees that developing a memorandum of understanding about the responsibilities of the State Board and CDE through mediation may be helpful in improving communications among all relevant parties. CDE staff has and continues to meet with the publisher and State Board on a weekly basis in order to keep the State Board apprised of the latest developments for the STAR program.
3. CDE staff are working on a plan with the test publisher for producing valid, reliable, and comparable individual student scores in reference to the STAR augmented tests. The spring 2000 results, as well as the field-test results, will be used to inform this work. Through meetings of representatives of the State Board, CDE, and the publisher the plan and timeline are being developed. The plan will include all the requirements set forth in Education Code section 60604.
4. CDE implemented weekly telephone conferences with the publisher once there was a contractual relationship. CDE uses weekly conferences with every publisher to review completed work, discuss and resolve any problems, and to address and plan the next implementation steps. Discussion of the next step(s) includes assignment of responsible staff. CDE will continue these valuable weekly conferences.
5. Current language in the CDE's contract with the publisher provides assurance that the test publisher has implemented most of the recommendations of the test publisher's

consultant. The 2001 contract (to be finalized by October 2000) will require the publisher to provide documentation that their consultant's recommendations relating to quality assurance procedures, subcontractor and contractor system development controls and management change procedures have been implemented.

6. CDE will calculate the costs of requiring all school district and school personnel to attend training sessions and will adopt regulations if there are no identified State-mandated costs.

⑩ Lastly, important changes have occurred on the membership of the State Board that have fundamentally changed the nature of communications with CDE. The draft report fails to mention this change in personnel or the change in personnel at the test publisher. We believe there has been and continues to be a dramatic improvement in communication and cooperation. If you have questions, please feel free to contact myself at 657-2644 or CDE's Audit Response Coordinator, Peggy Peters, at 657-4440.

Sincerely,

(Signed by: Scott Hill)

SCOTT HILL
Chief Deputy Superintendent of Public Instruction

COMMENTS

California State Auditor's Comments on the Response From the California Department of Education

To provide clarity and perspective, we are commenting on the response to our audit from the California Department of Education (department). The numbers below correspond to those we have placed in the department's response.

- ① The Joint Legislative Audit Committee asked us to conduct an audit regarding the implementation and execution of the STAR program. Therefore, our audit scope was not limited to just the 1998 and 1999 administrations.
- ② We acknowledge the efforts made by the superintendent of public instruction (superintendent) to recommend a single achievement test to the State Board of Education (board) on page 17 of our report. State law gave the board the sole discretion for designating the achievement test. The department fails to mention that this same state law also gave the board the option to request an evaluation of the achievement tests by independent experts.
- ③ The superintendent suggests that the 1,000 separate contracts between the test publisher and the school districts minimized her authority over the program. We disagree. Although the test publisher had a contractual relationship with each school district, as we state on page 19 of our report, the superintendent had a legal requirement, and therefore the authority, to establish a method for working with the test publisher to ensure that the test produced valid, reliable, and comparable results.
- ④ The department correctly states that it provided us with copies of correspondence from the superintendent to the board president regarding communication with the test publisher. Throughout the implementation of the STAR program there have been several instances of poor communication between the superintendent, department, and board. While we chose not to include this particular example, our report provides balance and objectivity by pointing out that all parties involved are guilty of poor communication.

- ⑤ The department has not demonstrated to us that sufficient internal planning has occurred for the STAR program. During our final exit conference on March 2, 2000, we again asked the department to provide us with its planning documents. We were given a department-wide strategic plan that contained a brief discussion of the STAR program but no mention of other factors, such as the department's criteria for measuring the reliability of test results or implementation timelines. Moreover, the department is attempting to minimize the superintendent's failure to prepare and submit to the Legislature an annual implementation plan, five-year cost projection, and timetable for the STAR program, as state law requires. The purpose of an annual implementation plan is to continually update long-range plans with new information and to make adjustments for significant changes to the program. Finally, the department is introducing information concerning the California Learning Assessment System, Pupil Testing Incentive Program (PTIP), and the High School Exit Exam that have no bearing on its responsibility for adequately planning the implementation of the STAR program.
- ⑥ While the department correctly states that the contract requires the test publisher to develop 24 writing questions, it fails to mention that the contract does not contain language that explicitly requires the test publisher to provide it with all 24. Nevertheless, our report focuses only on the miscommunication that occurred concerning the test publisher's fulfillment of its contractual obligation to field test 16 writing questions.
- ⑦ The department incorrectly states that a disconnect exists between our executive summary and report. Page 1 of our report clearly states that disputes between "these parties," which includes the board, caused the test publisher to halt production. This is consistent with page 14 of our report. In addition, the test publisher told us that it incurred significant expense by requiring its staff to work overtime to ensure that the field testing began on schedule.
- ⑧ Thus far, we have been the only independent objective party to study this series of events. We talked to all parties and reviewed relevant documents, and only then did we draw our conclusion. We do not believe that it is necessary to present each statement of fact made by the board, department, and test publisher concerning the initiation of the discussion with a board member about adding five math questions. More importantly, as we state on page 15 of our report, our point was to demonstrate the poor

communication that exists between the board, department, and test publisher. The department acknowledges this poor communication in its response.

- ⑨ The department would have the reader believe that it had no knowledge that the test publisher and board staff were proceeding with the board member's request to add five questions to the supplemental math test; however it misrepresents the facts. On January 5, 2000, the test publisher sent an e-mail to the department's Standards and Assessment Division stating that board staff asked it to increase the number of questions on the supplemental math test. The publisher outlined its approach for accomplishing its task. Also on January 5, 2000, board staff sent an e-mail to the deputy superintendent of the department's Accountability Branch stating that the test publisher had been asked to look into adding more questions to the supplemental math test.
- ⑩ The department's discussion adds little value for clarifying why a plan did not exist to ensure that the supplemental math test produced reliable results or how the necessary reliability will be achieved. Consequently, as we state on page 16 of our report, the board directed the test publisher to develop a plan to ensure the necessary reliability.
- ⑪ The department has missed the point. As we state on page 19 of our report, state law, specifically Education Code 60604(a)(2), requires the superintendent to establish a method of working with the test publisher to ensure that the achievement test results are valid, reliable, and comparable. This legal requirement gave the superintendent the authority to monitor the publisher's performance. Therefore, we are attributing the minimal monitoring of the test publisher's performance to the superintendent's failure to fulfill this legal requirement, not her failure to develop an implementation plan.
- ⑫ The department correctly states that it provided us with some information on its staffing requests. We reviewed this information and found that the reasons its requests were denied appeared reasonable. In the 1997-98 Budget Act, the governor eliminated \$1.7 million related to the department's PTIP in anticipation of the establishment of the STAR program. However, the governor also set aside \$800,000 for the establishment of the STAR program. Then, in August 1997, the superintendent requested \$505,230 in additional STAR funding primarily to support preparations for making a recommendation to the board. In the 1998-99 Budget

Act, the governor eliminated the 6.2 positions that the department requested, stating that since there were unfunded positions in its Standards and Assessments Division, there was no need for additional position authority. The department did not provide us with its staffing requests for subsequent years.

- ⑬ The department is incorrect. On March 28, 2000, the board told us that it has not adopted a policy stating that individual members should not engage in direct negotiations with the test publisher.
- ⑭ The text has been modified to reflect the department's concern.
- ⑮ As stated on page 25 of our report, we acknowledge that there is no legal requirement for the school districts to attend the department's or publisher's training sessions. However, we stand by our conclusion that school districts and testing personnel must take a greater role in ensuring the accuracy, completeness, and timeliness of the achievement test results. In fact in its response, the department indicates that it will address this issue.
- ⑯ The changes in board membership did not occur until February 25, 2000, and March 17, 2000, thus it is too early to conclude on the effect that these changes will have on the department's relationship with the board. Additionally, despite the test publisher's appointment of a California-based STAR program director in September 1999, as we demonstrate in our report, poor communication continues to exist.

cc: Members of the Legislature
Office of the Lieutenant Governor
Milton Marks Commission on California State
Government Organization and Economy
Department of Finance
Attorney General
State Controller
State Treasurer
Legislative Analyst
Senate Office of Research
California Research Bureau
Capitol Press