

November 2009

STUDENT ACHIEVEMENT

Schools Use Multiple
Strategies to Help
Students Meet
Academic Standards,
Especially Schools with
Higher Proportions of
Low-Income and
Minority Students



GAO

Accountability * Integrity * Reliability



Highlights of [GAO-10-18](#), a report to congressional committees

Why GAO Did This Study

The federal government has invested billions of dollars to improve student academic performance, and many schools, teachers, and researchers are trying to determine the most effective instructional practices with which to accomplish this. The Conference Report for the Consolidated Appropriations Act for Fiscal Year 2008 directed GAO to study strategies used to prepare students to meet state academic achievement standards. To do this, GAO answered: (1) What types of instructional practices are schools and teachers most frequently using to help students achieve state academic standards, and do those instructional practices differ by school characteristics? (2) What is known about how standards-based accountability systems have affected instructional practices? (3) What is known about instructional practices that are effective in improving student achievement? GAO analyzed data from a 2006-2007 national survey of principals and 2005-2006 survey of teachers in three states, conducted a literature review of the impact of standards-based accountability systems on instructional practices and of practices that are effective in improving student achievement, and interviewed experts.

What GAO Recommends

GAO makes no recommendations in this report. Education provided comments about issues pertaining to the study's approach that it believes should be considered. GAO clarified the report as appropriate.

[View GAO-10-18 or key components.](#)
For more information, contact Cornelia Ashby at (202) 512-7215 or AshbyC@gao.gov.

STUDENT ACHIEVEMENT

Schools Use Multiple Strategies to Help Students Meet Academic Standards, Especially Schools with Higher Proportions of Low-Income and Minority Students

What GAO Found

Nationwide, most principals focused on multiple strategies to help students meet academic standards, such as using student data to inform instruction and increasing professional development for teachers, according to our analysis of data from a U.S. Department of Education survey. Many of these strategies were used more often at high-poverty schools—those where 75 percent or more of the students were eligible for the free and reduced-price lunch program—and high-minority schools—those where 75 percent or more of students were identified as part of a minority population, than at lower poverty and minority schools. Likewise, math teachers in California, Georgia, and Pennsylvania increased their use of certain instructional practices in response to their state tests, such as focusing more on topics emphasized on assessments and searching for more effective teaching methods, and teachers at high-poverty and high-minority schools were more likely than teachers at lower-poverty schools and lower-minority schools to have made these changes, according to GAO's analysis of survey data collected by the RAND Corporation. Some researchers suggested that differences exist in the use of these practices because schools with lower poverty or lower minority student populations might generally be meeting accountability requirements and therefore would need to try these strategies less frequently.

Research shows that standards-based accountability systems can influence instructional practices in both positive and negative ways. For example, some research notes that using a standards-based curriculum that is aligned with corresponding instructional guidelines can facilitate the development of higher order thinking skills in students. But, in some cases, teacher practices did not always reflect the principles of standards-based instruction, and the difficulties in aligning practice with standards were attributed, in part, to current accountability requirements. Other research noted that assessments can be powerful tools for improving the learning process and evaluating student achievement, but assessments can also have some unintended negative consequences on instruction, including narrowing the curriculum to only material that is tested.

Many experts stated that methodological issues constrain knowing more definitively the specific instructional practices that improve student learning and achievement. Nevertheless, some studies and experts pointed to instructional practices that are considered to be effective in raising student achievement, such as differentiated instruction. Professional development for teachers was also highlighted as important for giving teachers the skills and knowledge necessary to implement effective teaching practices.

Contents

Letter		1
	Background	3
	Principals and Teachers Used a Variety of Instructional Practices to Help Students Meet Standards, and Many of These Practices Were Used More Frequently at Schools with Higher Proportions of Low-Income and Minority Students	5
	Research Shows That Standards-based Accountability Systems Can Influence Instructional Practices through Standards and Assessments in Both Positive and Negative Ways	15
	Research Highlights Some Potentially Successful Practices for Improving Student Achievement, although Experts Contend That Methodological Issues Constrain Reaching Definitive Conclusions about What Works	18
	Agency Comments and Our Evaluation	21
Appendix I	Scope and Methodology	26
Appendix II	Analyses of the Relationship between School Characteristics and Principals' Focus on School Improvement Strategies	29
Appendix III	List of Education Researchers	31
Appendix IV	Studies Meeting GAO's Criteria for Methodological Quality	32
Appendix V	Comments from the Department of Education	35
Appendix VI	GAO Contact and Staff Acknowledgments	37

Table

Table 1: Odds Ratios Indicating the Difference in Likelihood of Principals to Make School Improvement Strategies a Moderate or Major Focus after Controlling for Different Factors	30
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----

Figures

Figure 1: Principals' Responses Indicating That a School Improvement Strategy Was a Major or Moderate Focus of the School Improvement Efforts	7
Figure 2: Percent of Elementary and Middle School Math Teachers Who Reported Increasing Their Use of Certain Instructional Practices as a Result of State Test	11
Figure 3: How Survey Responses Differed between Math Teachers at High-Poverty and Low-Poverty Schools in Three States	13

Abbreviations

AYP	adequate yearly progress
ESEA	Elementary and Secondary Education Act of 1965
IASA	Improving America's Schools Act of 1994
NCLBA	No Child Left Behind Act of 2001
NLS-NCLB	National Longitudinal Study of No Child Left Behind
NSF	National Science Foundation
RAND	The RAND Corporation

This is a work of the U.S. government and is not subject to copyright protection in the United States. The published product may be reproduced and distributed in its entirety without further permission from GAO. However, because this work may contain copyrighted images or other material, permission from the copyright holder may be necessary if you wish to reproduce this material separately.



United States Government Accountability Office
Washington, DC 20548

November 16, 2009

Congressional Committees

The federal government has invested billions of dollars to help schools meet requirements of the No Child Left Behind Act of 2001 (NCLBA) to improve student academic performance in reading, math, and science.¹ To this end, many schools, teachers, and researchers are trying to determine the most effective instructional practices to improve student achievement. Instructional practices refer to school or district-level improvement strategies, such as aligning curriculum with academic standards, restructuring the school day, or providing additional professional development to teachers.² Instructional practices can also refer to classroom teaching practices like assigning more homework or searching for more effective teaching methods. Little is known about the extent to which instructional practices have changed in response to NCLBA's accountability requirements, whether these practices vary by type of school, and the extent to which some practices have proven to be more effective than others.

Under NCLBA, states are required to develop challenging student academic achievement standards, administer tests based on those standards (standards-based assessments) to measure student proficiency, and develop targets for performance on these tests. Specifically, NCLBA requires states to develop a plan to ensure that their students are making adequate yearly progress (AYP) toward proficiency in reading, math, and science by 2014 for students collectively and in key student subgroups, including low-income and minority students.

While NCLBA creates requirements for student proficiency, it generally allows states to determine how best to meet those requirements. The Conference Report accompanying the Consolidated Appropriations Act for Fiscal Year 2008 directed that GAO conduct a study of strategies used to prepare students to meet state academic achievement standards. In response, we agreed with the Senate and House Appropriations Committees, the Senate Committee on Health, Education, Labor and

¹Pub. L. No. 107-110.

²We use the phrase "instructional practices" to include tools for improving classroom teaching practices, such as providing additional professional development.

Pensions, and the House Committee on Education and Labor to address the following questions:

1. What types of instructional practices are schools and teachers most frequently using to help students achieve state academic standards, and do those instructional practices differ by school characteristics?
2. What is known about how standards-based accountability systems such as that in NCLBA have affected instructional practices?
3. What is known about instructional practices that are effective in improving student achievement?

To answer these questions, we analyzed data from two recent surveys of principals and teachers that were conducted by the RAND Corporation (RAND). The first survey, the nationally representative National Longitudinal Study of No Child Left Behind (NLS-NCLB), was sponsored by the U.S. Department of Education (Education) and asked principals the extent to which their schools were focusing on certain strategies in their school improvement efforts.³ We conducted an analysis of the school year 2006-2007 survey data on school improvement strategies by controlling for school characteristic variables, such as the percentage of a school's students receiving free or reduced price lunch (poverty); the percentage of students who are a racial minority (minority); whether the school is in an urban, urban fringe or large town, or rural area (school location); and the school's AYP performance status. The second survey, a three-state survey sponsored by the National Science Foundation (NSF), asked elementary and middle school teachers in California, Georgia, and Pennsylvania how their classroom teaching strategies differed due to a state math test.⁴

³State and Local Implementation of the No Child Left Behind Act Volume III—Accountability under NCLB: Interim Report. A report from the National Longitudinal Study of No Child Left Behind (NLS-NCLB) and the Study of State Implementation of Accountability and Teacher Quality under No Child Left Behind (SSI-NCLB) Kerstin Carlson Le Floch, AIR, Felipe Martinez, RAND, Jennifer O'Day, AIR, Brian Stecher, RAND, James Taylor, AIR, Andrea Cook, AIR. Prepared for: U.S. Department of Education Office of Planning, Evaluation and Policy Development Policy and Program Studies Service (2007).

⁴Laura S. Hamilton, Brian M. Stecher, Julie A. Marsh, Jennifer Sloan McCombs, Abby Robyn, Jennifer Lin Russell, Scott Naftel, and Heather Barney. "Standards-Based Accountability under No Child Left Behind: Experiences of Teachers and Administrators in Three States." Sponsored by the National Science Foundation. RAND 2007. The survey also asked about reported changes in strategies for science instruction as a result of the state science test, but we are only reporting on math instruction.

RAND selected these states to represent a range of approaches to standards-based accountability and to provide some geographic and demographic diversity. Using school year 2005-2006 data from the three-state survey, which is representative only of those three states individually, we measured associations between the teacher responses and the school characteristic variables. As part of these survey analyses, we reviewed documentation and performed electronic testing of the data obtained through the surveys and conducted interviews with the primary RAND researchers responsible for the data collection and analysis. We determined the survey data were sufficiently reliable for the purposes of our study. To answer questions two and three, we conducted a literature review and synthesis.⁵ We supplemented our synthesis by interviewing prominent education researchers identified in frequently cited articles and through discussions with other knowledgeable individuals.⁶ We also reviewed relevant federal laws and regulations.

We conducted our work from July 2008 to November 2009 in accordance with all sections of GAO's Quality Assurance Framework that are relevant to our objectives. The framework requires that we plan and perform the engagement to obtain sufficient and appropriate evidence to meet our stated objectives and to discuss any limitations in our work. We believe that the information and data obtained, and the analysis conducted, provide a reasonable basis for any findings and conclusions in this product.

Background

NCLBA reauthorized the Elementary and Secondary Education Act of 1965 (ESEA)⁷ and built upon accountability requirements created under a previous reauthorization, the Improving America's Schools Act of 1994 (IASA).⁸ Under ESEA, as amended, Congress sought to improve student

⁵Of the 20 studies we used that met our criteria for methodological quality, we relied heavily on two literature syntheses conducted by the Department of Education because of the large number of studies they included and the breadth of the topics they covered. For a list of these and the other studies meeting our criteria for methodological quality, see appendix IV. Additionally, a few other studies are cited in footnotes throughout the report but not included in the list of studies that we formally reviewed. Those cited in the footnotes were used because they provided more details or supplementary information about points that the experts made during our interviews.

⁶For a list of knowledgeable individuals with whom we spoke, see appendix III.

⁷Pub. L. No. 89-10.

⁸Pub. L. No. 103-382.

learning by incorporating academic standards and assessments in the requirements placed on states. Academic standards, which describe what students should know and be able to do at different grade levels in different subjects, help guide school systems in their choice of curriculum and help teachers plan for classroom instruction. Assessments, which states use to measure student progress in achieving the standards, are required to be administered by states.

NCLBA further strengthened some of the accountability requirements contained in ESEA, as amended. Specifically, NCLBA's accountability provisions require states to develop education plans that establish academic standards and performance goals for schools to meet AYP and lead to 100 percent of their students being proficient in reading, math, and science by 2014. This proficiency must be assessed annually in reading and math in grades 3 through 8 and periodically in science, whereas assessments were required less frequently under the IASA.⁹ Under NCLBA, schools' assessment data generally must be disaggregated to assess progress toward state proficiency targets for students in certain designated groups, including low-income students, minority students, students with disabilities, and those with limited English proficiency. Each of these groups must make AYP in order for the school to make AYP. Schools that fail to make AYP for 2 or more consecutive years are required to implement various improvement measures identified in NCLBA, and these measures are more extensive than those required under IASA. Education, which has responsibility for general oversight of NCLBA, reviews and approves state plans for meeting AYP requirements. As we have previously reported, Education had approved all states' plans—fully or conditionally—by June 2003.¹⁰

NCLBA also recognizes the role of teachers in providing a quality education by requiring states to ensure that all teachers in core academic subjects are "highly qualified." Under this requirement, teachers generally

⁹Assessments in science, which were first required under NCLBA in school year 2007-2008, are required at least once in grades 3 to 5, grades 6 to 9, and grades 10 to 12. High school students are required only to be assessed once in math and reading or language arts. In addition to annual assessments, high schools must include students' graduation rate, and elementary and middle schools must include one other academic indicator determined by the state to assess whether they made AYP.

¹⁰GAO, *No Child Left Behind Act: Improvements Needed in Education's Process for Tracking States' Implementation of Key Provisions*, [GAO-04-734](#) (Washington, D.C.: Sept. 30, 2004).

must have a bachelor's degree, be fully certified, and demonstrate their knowledge of the subjects they teach. Previously, there were no specific requirements regarding teacher quality under ESEA, as amended.¹¹

Principals and Teachers Used a Variety of Instructional Practices to Help Students Meet Standards, and Many of These Practices Were Used More Frequently at Schools with Higher Proportions of Low-Income and Minority Students

According to our analysis of NLS-NCLB data from Education, most principals reported their schools focused on multiple instructional practices in their voluntary school improvement efforts.¹² These strategies were used more often at schools with higher proportions of low-income students (“high-poverty schools”) and schools with higher proportions of minority students (“high-minority schools”) than at schools with lower proportions of low-income students (“low-poverty schools”) and schools with lower proportions of minority students (“low-minority schools”).¹³ Likewise, the survey of math teachers in California, Georgia, and Pennsylvania indicates teachers were using many different instructional practices in response to their state tests, and teachers at high-poverty and high-minority schools were more likely than teachers at low-poverty and low-minority schools to have been increasing their use of some of these practices. Some researchers we spoke with suggested that differences in the use of these instructional practices exist because schools with low-poverty or low-minority student populations might generally be meeting accountability standards and, therefore, would need to try these strategies less frequently.

¹¹For more information on teacher quality, see GAO, *Teacher Quality: Sustained Coordination among Key Federal Education Programs Could Enhance State Efforts to Improve Teacher Quality*, [GAO-09-593](#) (Washington, D.C.: July 2009).

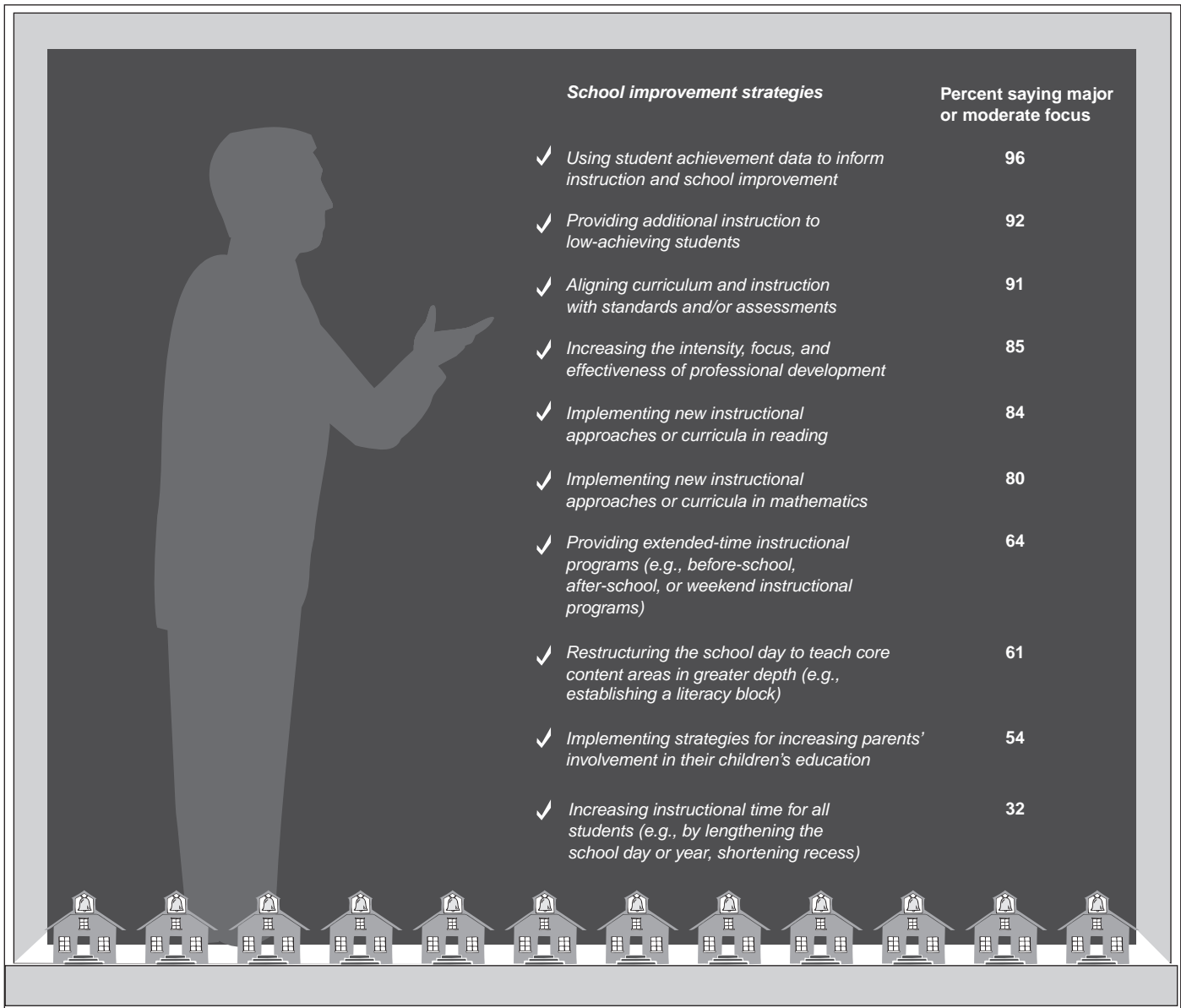
¹²For purposes of this report, we use the term “school improvement” to refer to the voluntary strategies used by school administrators and teachers to address various challenges within a school. By way of contrast, under NCLBA, schools that are identified for “school improvement” are those that have failed to make AYP for 2 or more consecutive years. These schools must implement certain activities identified in NCLBA that are meant to improve student academic achievement.

¹³Education classified schools as having “high—75 percent or more,” “moderate—35 to less than 75,” or “low—35 percent or less” percentages of low-income students using the number of students at the school that were eligible for the free and reduced-price lunch program. Schools were classified as having “high—75 percent or more,” “moderate—25 to less than 75,” or “low—25 percent or less” percentages of minority students, based on the school population that principals reported to be American Indian/Alaskan Native, Asian, Black or African-American, Hispanic or Latino, and Native Hawaiian or other Pacific Islander. Schools also were classified as central city (urban), urban fringe/large town (suburban), or small/fringe town (rural).

**Principals at High-Poverty
and High-Minority Schools
Emphasized Certain
School Improvement
Strategies More Than
Principals at Other Schools**

According to nationally representative data from Education's NLS-NCLB, in school year 2006-2007 most principals focused on multiple strategies in their school improvement efforts. The survey asked principals the extent to which their schools were focusing on ten different strategies in their voluntary school improvement initiatives. The three most common strategies were: (1) using student achievement data to inform instruction and school improvement; (2) providing additional instruction to low-achieving students; and (3) aligning curriculum and instruction with standards and/or assessments. (See fig. 1.) Nearly all school principals placed a major or moderate focus on three or more surveyed strategies in their school improvement efforts, and over 80 percent of principals placed a major or moderate focus on six or more strategies. However, as Education's report on the survey data cautioned, the number of improvement strategies emphasized was not necessarily an indication of the intensity or quality of the improvement efforts.

Figure 1: Principals' Responses Indicating That a School Improvement Strategy Was a Major or Moderate Focus of the School Improvement Efforts



Sources: GAO analysis of school year 2006-2007 NLS-NCLB survey data, Art Explosion (images).

Note: Some of the voluntary school improvement strategies identified above are similar to the corrective actions and restructuring options schools identified for improvement under NCLBA are required to choose from in preparing their school improvement plan. For example, implementing a new curriculum and extending the school day are both voluntary improvement strategies and possible strategies for improvement under the law.

While nearly all principals responded that they used multiple improvement strategies, there were statistically significant differences in principals' responses across a range of school characteristics, including percentage of the school's students receiving free or reduced price lunch (poverty), percentage of minority students, the school's location, and AYP status.¹⁴ For example, when comparing schools across poverty levels, we found that principals at high-poverty schools were two to three times more likely than principals at low-poverty schools to focus on five particular strategies in their school improvement efforts:

- Restructuring the school day to teach core content areas in greater depth;¹⁵
- Increasing instructional time for all students (e.g., by lengthening the school day or year, shortening recess);
- Providing extended-time instructional programs (e.g., before-school, after-school, or weekend instructional programs);
- Implementing strategies for increasing parents' involvement in their children's education; and
- Increasing the intensity, focus, and effectiveness of professional development.¹⁶

Likewise, when comparing schools across minority levels, we found that principals at high- and moderate-minority schools were approximately two to three times more likely than principals at low-minority schools to make six particular school improvement strategies a major or moderate focus of their school improvement efforts.¹⁷ For instance, principals at schools with a high percentage of minority students were more than three times as likely as principals at schools with a low percentage of minority students

¹⁴See appendix II for additional information about how principals' responses differed across school characteristics.

¹⁵Core content areas include those subjects for which testing is required under NCLBA—specifically, reading, math, and science.

¹⁶For the last three of these five strategies and one other—providing additional instruction to low-achieving students—there were also significant differences between moderate-poverty and low-poverty schools.

¹⁷See appendix II for a table that indicates which six strategies differed by school minority level.

to provide extended-time instruction such as after-school programs. A school's location was associated with differences in principals' responses about the strategies they used as well: principals at rural schools were only about one-third to one-half as likely as central city schools to make five of these school improvement strategies a moderate or major focus of their school improvement efforts.¹⁸

When we compared principal responses based on AYP status, there was some evidence of a statistically significant association between AYP status and the extent to which principals focused these strategies in their school improvement efforts, but it was limited when the other variables such as poverty and minority were taken into account. AYP status had some correlation with the demographic characteristics of poverty and minority, and those characteristics explained the patterns of principals' responses more fully than the AYP characteristic. However, our analysis generally showed that schools that had not made AYP were more likely to make six of these school improvement strategies a moderate or major focus of their school improvement plan than schools that had made AYP. Additionally, Education reported that schools identified for improvement under NCLBA—that is, schools that have not made AYP for two or more consecutive years—were engaged in a greater number of improvement efforts than non-identified schools. Therefore, principals of the non-identified schools may have been less likely than principals of identified schools to view specific strategies as a major or moderate focus.

We spoke with several researchers about the results of our analysis of the principals' responses, especially at high-poverty and high-minority schools. While the researchers could not say with certainty the reasons for the patterns, they noted that high-poverty and high-minority schools tend to be most at risk of not meeting their states' standards, so that principals at those schools might be more willing to try different approaches. Conversely, the researchers noted that principals at schools meeting

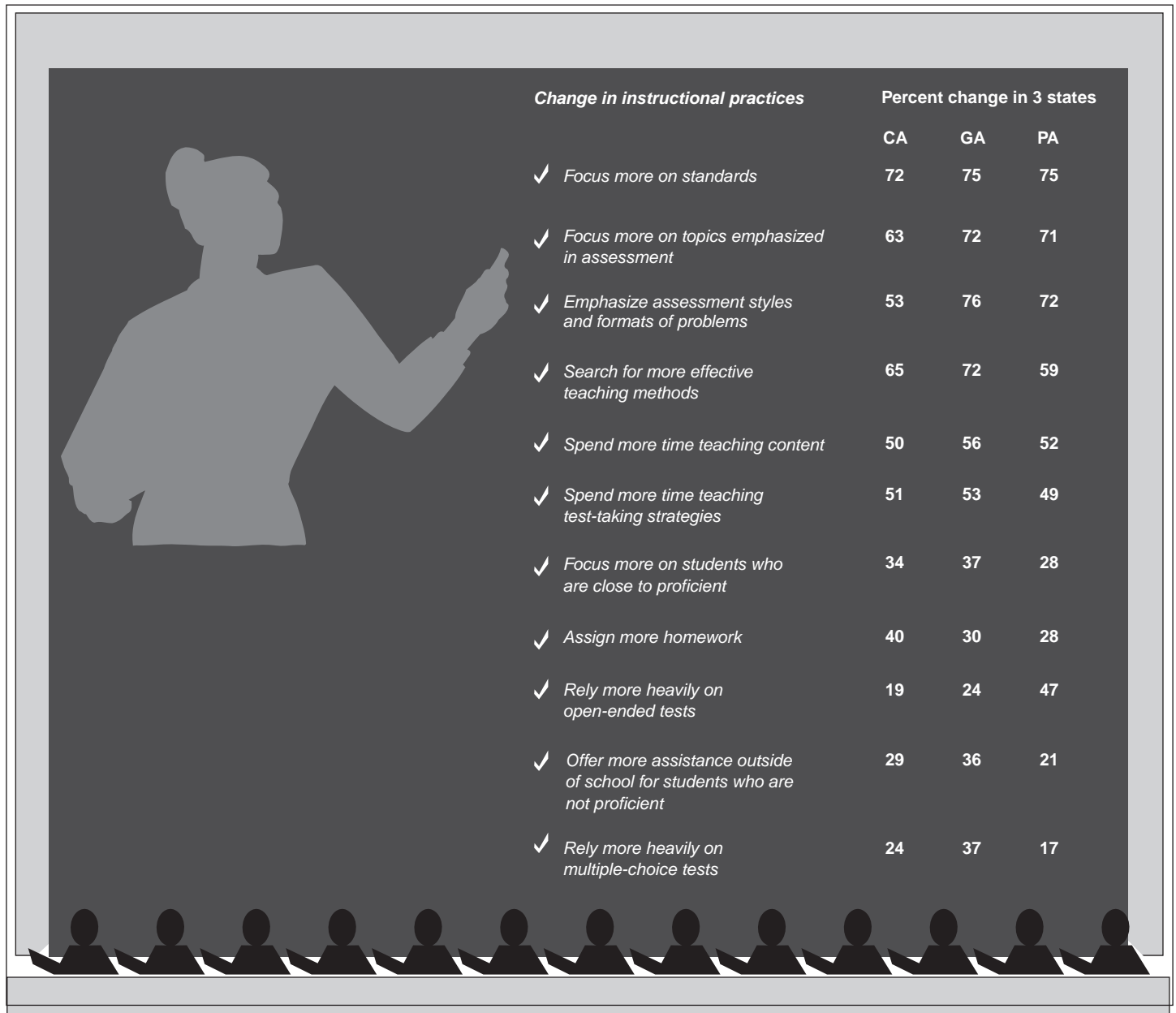
¹⁸Urban fringe or large town schools were no different from the central city schools with respect to making these strategies a major or moderate focus. In the 2003-2004 school year, about 30 percent of all U.S. elementary and secondary public schools were located in rural areas and approximately 20 percent of public school students were enrolled in rural schools. See S. Provasnik, A. KewalRamani, M. M. Coleman, L. Gilbertson, W. Herring, and Q. Xie, *Status of Education in Rural America* (NCES 2007-040). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education (Washington, D.C.: 2007). See appendix II for a table that indicates which five strategies differed by school geographic type.

standards would not have the same incentives to adopt as many school improvement strategies.

Most Math Teachers in Three Surveyed States Have Increased Their Use of Certain Instructional Practices in Response to State Tests, especially in High-Poverty and High-Minority Schools

The RAND survey of elementary and middle school math teachers in California, Georgia and Pennsylvania showed that in each of the three states at least half of the teachers reported increasing their use of certain instructional practices in at least five areas as a result of the statewide math test (see fig. 2). For example, most teachers in Pennsylvania responded that due to the state math test they: (1) focused more on standards, (2) emphasized assessment styles and formats, (3) focused more on subjects tested, (4) searched for more effective teaching methods, and (5) spent more time teaching content.

Figure 2: Percent of Elementary and Middle School Math Teachers Who Reported Increasing Their Use of Certain Instructional Practices as a Result of State Test



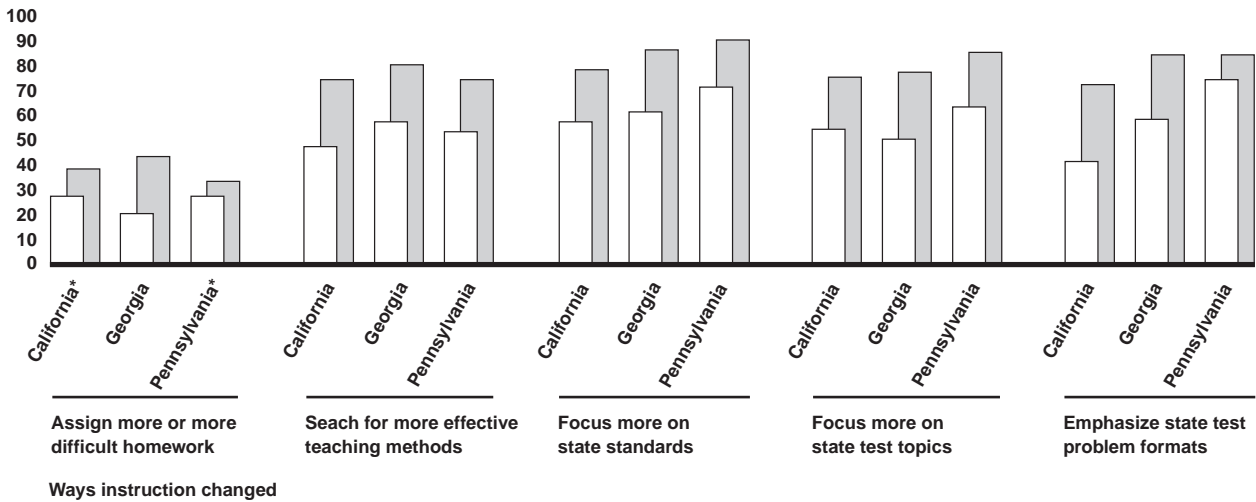
Sources: GAO analysis of 2005 survey data from Standards-Based Accountability Under No Child Left Behind: Experiences of Teachers and Administrators in Three States. Hamilton et al. Art Explosion (images).

As we did with the survey responses of principals, we analyzed the teacher survey data to determine whether math teachers' responses differed by school characteristics for poverty, minority, location, and AYP status. As with the principals' responses, we found that elementary and middle school math teachers in high-poverty and high-minority schools were more likely than teachers in low-poverty and low-minority schools to report increasing their use of certain instructional practices, and this pattern was consistent across the three states (see fig. 3). For example, 69 percent of math teachers at high-poverty schools in California indicated they spent more time teaching test-taking strategies as opposed to 38 percent of math teachers in low-poverty schools. In Georgia, 50 percent of math teachers in high-poverty schools reported offering more outside assistance to non-proficient students in contrast to 26 percent of math teachers in low-poverty schools. Fifty-one percent of math teachers at high-poverty schools in Pennsylvania reported focusing more attention on students close to proficiency compared to 23 percent of math teachers doing so in low poverty schools.¹⁹

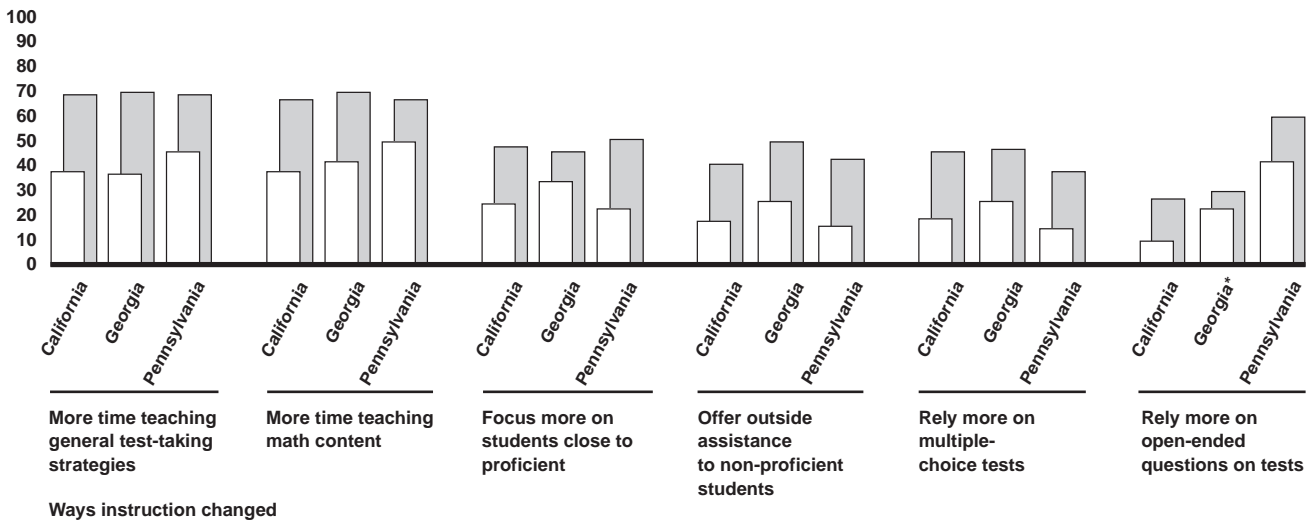
¹⁹When we compared moderate-poverty schools to high-poverty and low-poverty schools, we saw fewer statistically significant differences than in our high-poverty and low-poverty school comparison.

Figure 3: How Survey Responses Differed between Math Teachers at High-Poverty and Low-Poverty Schools in Three States

Percentage of teachers changing instructional practices as a result of state math test



Percentage of teachers changing instructional practices as a result of state math test



* = Not statistically significant with a 95 percent level of confidence for the difference between high-poverty and low-poverty schools

Low poverty
 High poverty

Sources: GAO analysis of school year 2005-2006 survey data from Standards-Based Accountability Under No Child Left Behind: Experiences of Teachers and Administrators in Three States. Hamilton et al.

Similar to what our poverty analysis showed, survey responses provided some evidence that math teachers in high-minority schools were more likely than those in low-minority schools to change their instructional practices. Math teachers at high-minority schools in each of the three states, as compared to those at low-minority schools, were more likely to:

- rely on open-ended tests in their own classroom assessments;
- increase the amount of time spent teaching mathematics by replacing non-instructional activities with mathematics instruction;
- focus on topics emphasized in the state math test; and
- teach general test-taking strategies.

We also analyzed the RAND data with regard to school location and a school's AYP status, but results from these characteristics were not significant for as many instructional practices.²⁰

As we did regarding the survey responses of principals, we spoke to several researchers, including the authors of the three-state teacher study, regarding possible reasons for the patterns we saw in the teacher survey data. The researchers we spoke with provided similar possible reasons for the patterns in the teacher survey as they did for patterns in the principal survey. For instance, the researchers noted that high-poverty and high-minority schools are more likely to be at risk of failing to meet the state standards, which might prompt teachers to try different approaches. On the other hand, the researchers stated that teachers at those schools meeting the standards would not have the same incentives to change their instructional practices.

²⁰For the three state data, we conducted a simple analysis that did not control for multiple factors, since we had access only to RAND's bi-variate analyses of the data rather than the data itself. Because of this, we could not perform a multivariate analysis, which would allow us to control for other factors.

Research Shows That Standards-based Accountability Systems Can Influence Instructional Practices through Standards and Assessments in Both Positive and Negative Ways

Research shows that using a standards-based curriculum that is aligned with corresponding instructional guidelines can positively influence teaching practices. Specifically, some studies reported changes by teachers who facilitated their students developing higher-order thinking skills, such as interpreting meaning, understanding implied reasoning, and developing conceptual knowledge, through practices such as multiple answer problem solving, less lecture and more small group work. Additionally, a few researchers we interviewed stated that a positive effect of NCLBA's accountability provisions has been a renewed focus on standards and curriculum.²¹

However, some studies indicated that teachers' practices did not always reflect the principles of standards-based instruction and that current accountability policies help contribute to the difficulty in aligning practice with standards. Some research shows that, while teachers may be changing their instructional practices in response to standards-based reform, these changes may not be fully aligned with the principles of the reform. That research also notes that the reliability in implementing standards in the classroom varied in accordance with teachers' different beliefs in and support for standards-based reform as well as the limitations in their instructional capabilities. For example, one observational study of math teachers showed that, while teachers implemented practices envisioned by standards-based reform, such as getting students to work in small groups or using manipulatives (e.g., cubes or tiles), their approaches did not go far enough in that students were not engaged in conversations about mathematical or scientific concepts and ideas.²² To overcome these challenges, studies point to the need for teachers to have opportunities to learn, practice, and reflect on instructional practices that incorporate the standards, and then to observe their effects on student learning. However, some researchers have raised concerns that current accountability systems' focus on test scores and mandated timelines for achieving proficiency levels for students do not give teachers enough time to learn, practice, and reflect on instructional practices and may discourage some

²¹The National Governors' Association and the Council of Chief State School Officers are coordinating a committee of experts to develop common academic standards for math and language arts skills. As of June 2009, 46 states had signed onto this effort to adopt the common standards once they were completed.

²²W. Firestone, R. Schorr, and L. Monfils, editors. *Ambiguity of Teaching to the Test: Standards, Assessments, and Educational Reform*, 160-161 (2004).

teachers from trying ambitious teaching practices envisioned by standards-based reform.

Another key element of a standards-based accountability system is assessments, which help measure the extent to which schools are improving student learning through assessing student performance against the standards. Some researchers note that assessments are powerful tools for managing and improving the learning process by providing information for monitoring student progress, making instructional decisions, evaluating student achievement, and evaluating programs. In addition, assessments can also influence instructional content and help teachers use or adjust specific classroom practices. As one synthesis concluded, assessments can influence whether teachers broaden or narrow the curriculum, focus on concepts and problem solving—or emphasize test preparation over subject matter content.²³

In contrast, some of the research and a few experts we interviewed raised concerns about testing formats that do not encourage challenging teaching practices and instructional practices that narrow the curriculum as a result of current assessment practices.²⁴ For example, depending on the test used, research has shown that teachers may be influenced to use teaching approaches that reflect the skills and knowledge to be tested. Multiple choice tests tend to focus on recognizing facts and information while open-ended formats are more likely to require students to apply critical thinking skills. Conclusions from a literature synthesis conducted by the Department of Education stated that “teachers respond to assessment formats used, so testing programs must be designed and administered with this influence in mind. Tests that emphasize inquiry, provide extended writing opportunities, and use open-ended response formats or a portfolio approach tend to influence instruction in ways quite different from tests that use closed-ended response formats and which emphasize

²³Helen S. Apthorp, et al., “Standards in Classroom Practice Research Synthesis,” Mid-Continent Research for Education and Learning (October 2001).

²⁴NCLBA added to the assessment requirements included in IASA. For example, NCLBA requires states to implement annual assessments for all students in every grade for grades 3-8 in reading and math; IASA required assessments at least once in each of three grade spans: 3-5, 6-9, and 10-12. Additionally, unlike IASA, NCLBA sets a uniform timeline for when all students must meet state proficiency targets.

procedures.”²⁵ We recently reported that states have most often chosen multiple choice items over other item types of assessments because they are cost effective and can be scored within tight time frames. While multiple choice tests provide cost and time saving benefits to states, the use of multiple choice items make it difficult, if not impossible, to measure highly complex content.²⁶ Other research has raised concerns that, to avoid potential consequences from low-scoring assessment results under NCLBA, teachers are narrowing the curriculum being taught—sometimes referred to as “teaching to the test”—either by spending more classroom time on tested subjects at the expense of other non-tested subjects, restricting the breadth of content covered to focus only on the content covered by the test, or focusing more time on test-taking strategies than on subject content.²⁷

²⁵P. A. Lauer, D. Snow, M. Martin-Glenn, R.J.Van Buhler, K. Stoutemyer, R. Snow-Renner, *The Influence of Standards on K-12 Teaching and Student Learning: A Research Synthesis*, Regional Education Laboratory, August 19, 2005, p. 91.

²⁶GAO, *No Child Left Behind Act: Enhancements in the Department of Education’s Review Process Could Improve State Academic Assessments*, [GAO-09-911](#), (Washington, D.C.: September 2009).

²⁷For example, according to data from Education’s national survey, about 18 percent of elementary school teachers reported that instruction time for math increased from school years 2004-2005 to 2006-2007, and about 22 percent of elementary school teachers reported that instruction time for reading/language arts increased over the same period. However, approximately three-quarters of teachers reported no change in instructional time in these two subjects. GAO, *Access to Arts Education: Inclusion of Additional Questions in Education’s Planned Research Would Help Explain Why Instruction Time Has Decreased for Some Students*, [GAO-09-286](#) (Washington, D.C.: February 2009). In addition, a report by the Department of Education states that from 1987-1988 to 2003-2004, teacher survey results from the Schools and Staffing Survey conducted by the National Center for Education Statistics indicate that elementary teachers had increased instructional time on reading and mathematics and decreased the amount of time spent on science and social studies during this period. See U.S. Department of Education, Office of Planning, Evaluation, and Policy Development, Policy and Program Studies Service, *Title I Implementation—Update on Recent Evaluation Findings* (Washington, D.C.: 2009).

Research Highlights Some Potentially Successful Practices for Improving Student Achievement, although Experts Contend That Methodological Issues Constrain Reaching Definitive Conclusions about What Works

Our literature review found some studies that pointed to instructional practices that appear to be effective in raising student achievement. But, in discussing the broader implications of these studies with the experts that we interviewed, many commented that, taken overall, the research is not conclusive about which specific instructional practices improve student learning and achievement.

Some researchers stated that this was due to methodological issues in conducting the research. For example, one researcher explained that, while smaller research studies on very specific strategies in reading and math have sometimes shown powerful relationships between the strategy used and positive changes in student achievement, results from meta-analyses of smaller studies have been inconclusive in pointing to similar patterns in the aggregate. A few other researchers stated that the lack of empirical data about how instruction unfolds in the classroom hampers the understanding about what works in raising student performance.

A few researchers also noted that conducting research in a way that would yield more conclusive results is difficult. One of the main difficulties, as explained by one researcher, is the number of variables a study may need to examine or control for in order to understand the effectiveness of a particular strategy, especially given the number of interactions these variables could have with each other. One researcher mentioned cost as a challenge when attempting to gather empirical data at the classroom level, stating “teaching takes place in the classroom, but the expense of conducting classroom-specific evaluations is a serious barrier to collecting this type of data.” Finally, even when research supports the efficacy of a strategy, it may not work with different students or under varying conditions. In raising this point, one researcher stated that “educating a child is not like making a car” whereby a production process is developed and can simply be repeated again and again. Each child learns differently, creating a challenge for teachers in determining the instructional practices that will work best for each student.

Some of the practices identified by both the studies and a few experts as those with potential for improving student achievement were:

- *Differentiated instruction.* In this type of instruction, teaching practices and plans are adjusted to accommodate each student’s skill level for the task at hand. Differentiated instruction requires teachers to be flexible in their teaching approach by adjusting the curriculum and presentation of information for students, thereby providing multiple options for students to take in and process information. As one researcher described it,

effective teachers understand the strategies and practices that work for each student and in this way can move all students forward in their learning and achievement.

- *More guiding, less telling.* Researchers have identified two general approaches to teaching: didactic and interactive. Didactic instruction relies more on lecturing and demonstrations, asking short answer questions, and assessing whether answers are correct. Interactive instruction focuses more on listening and guiding students, asking questions with more than one correct answer, and giving students choices during learning. As one researcher explained, both teaching approaches are important, but some research has shown that giving students more guidance and less direction helps students become critical and independent thinkers, learn how to work independently, and assess several potential solutions and apply the best one. These kinds of learning processes are important for higher-order thinking. However, implementing “less instruction” techniques requires a high level of skill and creativity on the part of the teacher.²⁸
- *Promoting effective discourse.* An important corollary to the teacher practice of guiding students versus directing them is effective classroom discussion. Research highlights the importance of developing students’ understanding not only of the basic concepts of a subject, but higher-order thinking and skills as well. To help students achieve understanding, it is necessary to have effective classroom discussion in which students test and revise their ideas, and elaborate on and clarify their thinking. In guiding students to an effective classroom discussion, teachers must ask engaging and challenging questions, be able to get all students to participate, and know when to provide information or allow students to discover it for themselves.

Additionally, one synthesis of several experimental studies examining practices in elementary math classrooms identified two instructional approaches that showed positive effects on student learning. The first was cooperative learning in which students work in pairs or small teams and

²⁸The final report of the National Mathematics Advisory Panel takes a slightly different position regarding this practice stating that “All-encompassing recommendations that instruction should be entirely ‘student centered’ or ‘teacher directed’ are not supported by research . . . High-quality research does not support the exclusive use of either approach.” National Mathematics Advisory Panel. *Foundations for Success: The Final Report of the National Mathematics Advisory Panel*, U.S. Department of Education (Washington, D.C.: 2008).

are rewarded based on how well the group learns. The other approach included programs that helped teachers introduce math concepts and improve skills in classroom management, time management, and motivation. This analysis also found that using computer-assisted instruction had moderate to substantial effects on student learning, although this type of instruction was always supplementary to other approaches or programs being used.

We found through our literature review and interviews with researchers that the issue of effective instructional practices is intertwined with professional development. To enable all students to achieve the high standards of learning envisioned by standards-based accountability systems, teachers need extensive skills and knowledge in order to use effective teaching practices in the classroom. Given this, professional development is critical to supporting teachers' learning of new skills and their application. Specifically, the research concludes that professional development will more likely have positive impacts on both teacher learning and student achievement if it:

- Focuses on a content area with direct links to the curriculum;
- Challenges teachers intellectually through reflection and critical problem solving;
- Aligns with goals and standards for student learning;
- Lasts long enough so that teachers can practice and revise their techniques;
- Occurs collaboratively within a teacher learning community—ongoing teams of teachers that meet regularly for the purposes of learning, joint lesson planning, and problem solving;
- Involves all the teachers within a school or department;
- Provides active learning opportunities with direct applications to the classroom; and
- Is based on teachers' input regarding their learning needs.

Some researchers have raised concerns about the quality and intensity of professional development currently received by many teachers nationwide. One researcher summarized these issues by stating that

professional development training for teachers is often too short, provides no classroom follow up, and models more “telling than guiding” practices. Given the decentralized nature of the U.S. education system, the support and opportunity for professional development services for teachers varies among states and school districts, and there are notable examples of states that have focused resources on various aspects of professional development. Nevertheless, shortcomings in teachers’ professional development experiences overall are especially evident when compared to professional development requirements for teachers in countries whose students perform well on international tests, such as the Trends in International Mathematics and Science Study and the Program for International Student Assessment. For example, one study showed that fewer than 10 percent of U.S. math teachers in school year 2003-04 experienced more than 24 hours of professional development in mathematics content or pedagogy during the year; conversely, teachers in Sweden, Singapore, and the Netherlands are required to complete 100 hours of professional development per year.²⁹

Agency Comments and Our Evaluation

We provided a copy of our draft report to the Secretary of Education for review and comment. Education’s written comments, which are contained in appendix V, expressed support for the important questions that the report addresses and noted that the American Recovery and Reinvestment Act of 2009 included \$250 million to improve assessment and accountability systems. The department specifically stated that the money is for statewide data systems to provide information on individual student outcomes that could help enable schools to strengthen instructional practices and improve student achievement. However, the department raised several issues about the report’s approach. Specifically, the department commented that we (1) did not provide the specific research citations throughout the report for each of our findings or clearly explain how we selected our studies; (2) mixed the opinions of education experts with our findings gleaned from the review of the literature; (3) did not present data on the extent to which test formats had changed or on the relationship between test format and teaching practices when discussing our assessment findings; and (4) did not provide complete information

²⁹L. Darling-Hammond, R. Wei, A. Andree, N. Richardson, and S. Orphanos, *Professional Learning in the Learning Profession: A Status Report on Teacher Development in the United States and Abroad, Technical Report* (National Staff Development Council and The School Redesign Network at Stanford University: February 2009) 18 and 22.

from an Education survey regarding increases and decreases in instructional time.

As stated in the beginning of our report, the list of studies we reviewed and used for our findings are contained in appendix IV. We provide a description in appendix I of our criteria, the types of databases searched, the types of studies examined (e.g., experimental and nonexperimental) and the process by which we evaluated them. We relied heavily on two literature syntheses conducted by the Department of Education—*Standards in Classroom Practice: Research Synthesis* and *The Influence of Standards on K-12 Teaching and Student Learning: A Research Synthesis*, which are included in the list. These two syntheses covered, in a more comprehensive way than many of the other studies that we reviewed, the breadth of the topics that we were interested in and included numerous research studies in their reviews. Many of the findings in this report about the research are taken from the conclusions reached in these syntheses. However, to make this fact clearer and more prominent, we added this explanation to our abbreviated scope and methodology section on page 5 of the report.

Regarding the use of expert opinion, we determined that obtaining the views of experts about the research we were reviewing would be critical to our understanding its broader implications. This was particularly important given the breadth and scope of our objectives. The experts we interviewed, whose names and affiliations are listed in appendix III, are prominent researchers who conduct, review, and reflect on the current research in the field, and whose work is included in some of the studies we reviewed, including the two literature syntheses written by the Department of Education and used by us in this study. We did not consider their opinions “conjecture” but grounded in and informed by their many years of respected work on the topic. We have been clear in the report as to when we are citing expert opinion, the research studies, or both.

Regarding the report section discussing the research on assessments, it was our intent to highlight that, according to the research, assessments have both positive and negative influences on classroom teaching practices, not to conclude that NCLBA was the cause of either. Our findings in this section of the report are, in large part, based on conclusions from the department’s syntheses mentioned earlier. For example, *The Influence of Standards on K-12 Teaching and Student Learning: A Research Synthesis* states “... tests matter—the content covered, the format used, and the application of their results—all influence teacher behavior.” Furthermore, we previously reported that

states most often have chosen multiple choice assessments over other types because they can be scored inexpensively and their scores can be released prior to the next school year as required by NCLBA.³⁰ That report also notes that state officials and alignment experts said that multiple choice assessments have limited the content of what can be tested, stating that highly complex content is “difficult if not impossible to include with multiple choice items.” However, we have revised this paragraph to clarify our point and provide additional information.

Concerning the topic of narrowing the curriculum, we agree with the Department of Education that this report should include a fuller description of the data results from the cited Education survey in order to help the reader put the data in an appropriate context. Hence, we have added information to that section of the report. However, one limitation of the survey data we cite is that it covers changes in instructional time for a short time period—from school year 2004-05 to 2006-07. In the its technical comments, the Department refers to its recent report, *Title I Implementation: Update on Recent Evaluation Findings* for a fuller discussion of this issue. The Title I report, while noting that most elementary teachers reported no change from 2004-05 to 2006-07 in the amount of instructional time that they spent on various subjects, also provides data over a longer, albeit earlier period time period, from 1987-88 to 2003-04, from the National Center on Education Statistics Schools and Staffing Survey. In analyzing this data, the report states that elementary teachers had increased instructional time on reading and mathematics and decreased the amount of time spent on science and social studies during this period. We have added this information as well. Taken together, we believe these data further reinforce our point that assessments under current accountability systems can have, in addition to positive influences on teaching, some negative ones as well, such as the curriculum changes noted in the report, even if the extent of these changes is not fully known.

Education also provided technical comments that we incorporated as appropriate.

³⁰GAO, *No Child Left Behind Act: Enhancements in the Department of Education’s Review Process Could Improve State Academic Assessments*, [GAO-09-911](#) (Washington, D.C.: September 2009).

We are sending copies of this report to the Secretary of Education, relevant congressional committees, and other interested parties. The report also is available at no charge on the GAO Web site at <http://www.gao.gov>.

If you or your staff have any questions about this report, please contact me at (202) 512-7215 or ashbyc@gao.gov. Contact points for our Office of Congressional Relations and Public Affairs may be found on the last page of this report. GAO staff who made major contributions to this report are listed in appendix VI.



Cornelia M. Ashby
Director, Education, Workforce, and
Income Security Issues

List of Congressional Committees

The Honorable Tom Harkin
Chairman
The Honorable Thad Cochran
Ranking Member
Subcommittee on Labor, Health and
Human Services, Education
and Related Agencies
Committee on Appropriations
United States Senate

The Honorable Dave Obey
Chairman
The Honorable Todd Tiahrt
Ranking Member
Subcommittee on Labor, Health and
Human Services, Education
and Related Agencies
Committee on Appropriations
House of Representatives

Appendix I: Scope and Methodology

To address the objectives of this study, we used a variety of methods. To determine the types of instructional practices schools and teachers are using to help students achieve state academic standards and whether those practices differ by school characteristics, we used two recent surveys of principals and teachers. The first survey, a nationally-representative survey from the Department of Education's (Education) National Longitudinal Study of No Child Left Behind (NLS-NCLB) conducted by the RAND Corporation (RAND), asked principals the extent to which their schools were focusing on certain strategies in their voluntary school improvement efforts. Education's *State and Local Implementation of the No Child Left Behind Act Volume III—Accountability Under NCLB: Interim Report* included information about the strategies emphasized by principals as a whole, and we obtained from Education the NLS-NCLB database to determine the extent to which principals' responses differed by school characteristic variables. We conducted this analysis on school year 2006-2007 data by controlling for four school characteristic variables: (1) the percentage of a school's students receiving free or reduced price lunch (poverty); (2) the percentage of students who are a racial minority (minority); (3) whether the school is in an urban, urban fringe (suburban), or rural area (school location); and (4) the school's adequate yearly performance (AYP) status.

We analyzed data from a second RAND survey, which was a three-state survey sponsored by the National Science Foundation that asked math teachers in California, Georgia, and Pennsylvania how their classroom teaching strategies differed due to a state math test.¹ RAND selected these states to represent a range of approaches to standards-based accountability and to provide some geographic and demographic diversity; the survey data is representative only for those three states individually. RAND's report on the three-state survey data included information about how teachers within each of the three states had changed their teaching practices due to a state accountability test.² RAND provided us with descriptive data tables based on its school year 2005-2006 survey data; we analyzed the data to measure associations between the strategies used and

¹Several education experts we spoke to said the list of practices was fairly complete, but one expert noted that professional development is also an important instructional practice.

²Laura S. Hamilton, Brian M. Stecher, Julie A. Marsh, Jennifer Sloan McCombs, Abby Robyn, Jennifer Lin Russell, Scott Naftel, and Heather Barney, "Standards-Based Accountability under No Child Left Behind: Experiences of Teachers and Administrators in Three States" (Sponsored by the National Science Foundation. RAND 2007).

the school characteristic variables.³ We requested tables that showed this information for teachers in all schools, and separately for teachers in different categories of schools (elementary and middle schools) and by the school characteristics of poverty, minority, school location and AYP status. We obtained from RAND standard error information associated with the estimates from the different types of schools and thus were able to test the statistical significance of differences in likelihood between what teachers from different types of schools reported.

As part of our analyses for both surveys, we reviewed documentation and performed electronic testing of the data obtained through the surveys. We also conducted several interviews with several researchers responsible for the data collection and analyses and obtained information about the measures they took to ensure data reliability. On the basis of our efforts to determine the reliability of the data, we determined the data from each of these surveys were sufficiently reliable for the purposes of our study.

We reviewed existing literature to determine what researchers have found regarding the effect of standards-based accountability systems on instructional practices, and practices that work in raising student achievement. To identify existing studies, we conducted searches of various databases, such as the Education Resources Information Center, Proquest, Dialog EDUCAT, and Education Abstracts. We also asked all of the education researchers that we interviewed to recommend additional studies. From these sources, we identified 251 studies that were relevant to our study objectives about the effect of standards-based accountability systems on instructional practices and instructional practices there are effective in raising student achievement. We selected them according to the following criteria: covered the years 2001 through 2008 and were either experimental or quasi-experimental studies, literature syntheses, or studied multiple sites.⁴ We selected the studies for our review based on their methodological strength, given the limitations of the methods used, and not necessarily on whether the results could be generalized. We performed our searches from August 2008 to January 2009.

³Scott Naftel, Laura S. Hamilton, and Brian M. Stecher, "Working Paper Supplemental Analyses of ISBA Survey Responses" (WR-628-EDU. RAND. November 2008).

⁴Some research, including the syntheses that we reviewed, included some studies outside these date parameters. Additionally, the syntheses used to support some of the findings were not meta-analyses but literature reviews, although both qualitative and quantitative studies were included in the syntheses.

To assess the methodological quality of the selected studies, we developed a data collection instrument to obtain information systematically about each study being evaluated and about the features of the evaluation methodology. We based our data collection and assessments on generally accepted social science standards. We examined factors related to the use of comparison and control groups; the appropriateness of sampling and data collection methods; and for syntheses, the process and criteria used to identify studies. A senior social scientist with training and experience in evaluation research and methodology read and coded the methodological discussion for each evaluation. A second senior social scientist reviewed each completed data collection instrument and the relevant documentation to verify the accuracy of every coded item. This review identified 20 selected studies that met GAO's criteria for methodological quality.

We supplemented our synthesis by interviewing prominent education researchers identified in frequently cited articles and through discussions with knowledgeable individuals. We also conducted interviews with officials at the U.S. Department of Education, including the Center on Innovation and Improvement, and the Institute on Education Sciences' National Center for Education Evaluation and Regional Assistance, as well as other educational organizations. We also reviewed relevant federal laws and regulations.

Appendix II: Analyses of the Relationship between School Characteristics and Principals' Focus on School Improvement Strategies

In order to analyze the National Longitudinal Study of No Child Left Behind (NLS-NCLB) principal survey conducted by the RAND Corporation, we analyzed strategies on which principals most often focused, taking into account the percentage of a school's students receiving free or reduced price lunch (poverty), the percentage of students who are a racial minority (minority), whether the school is in an urban, suburban, or rural area (school location), and the school's adequate yearly performance (AYP) status (see table 1).¹ Our analyses used "odds ratios," generally defined as the ratio of the odds of an event occurring in one group compared to the odds of it occurring in another group, to express differences in the likelihoods of schools with different characteristics using these strategies. We used odds ratios rather than percentages because they are more appropriate for statistical modeling and multivariate analysis. Odds ratios indicate how much higher (when they are greater than 1.0) or lower (when they are less than 1.0) the odds were that principals would respond that a given strategy was a major or moderate focus. We included a reference category for the school characteristics (low minority, low poverty, and central city) in the top row of table 1, and put comparison groups beneath those reference categories, as indicated by the column heading in the second row (high-minority, high-poverty, or rural schools). As an example, the third cell in the "high-minority schools" column indicates that principals in high-minority schools were 2.65 times more likely to make "implementing new instructional approaches or curricula in reading/language arts/English" a focus of their school improvement efforts. In another example, the odds that principals would "restructure the school day to teach core content areas in greater depth (e.g., establishing a literacy block)" were 2.8 times higher for high-poverty schools than low poverty schools, as seen in the sixth cell under "high-poverty schools." Those cells with an asterisk indicate statistically significant results; that is, we have a high degree of confidence that the differences we see are not just due to chance but show an actual difference in the survey responses. See appendix I for further explanation of our methodology.

¹Table 1 does not include AYP status, because we found that the demographic characteristics of poverty and minority explained the patterns of principals' responses more fully than AYP status.

**Appendix II: Analyses of the Relationship
between School Characteristics and
Principals' Focus on School Improvement
Strategies**

Table 1: Odds Ratios Indicating the Difference in Likelihood of Principals to Make School Improvement Strategies a Moderate or Major Focus after Controlling for Different Factors

School demographic School Improvement Strategy	(Compared to low-minority schools)		(Compared to low-poverty schools)		(Compared to central city schools)	
	High-minority schools	Middle minority schools	High-poverty schools	Middle poverty schools	Rural schools	Suburban/fringe schools
Using student achievement data to inform instruction and school improvement	1.24	3.01*	2.51	1.34	0.46	0.98
Aligning curriculum and instruction with standards and/or assessments	1.24	2.09*	1.81	0.92	0.58	0.79
Implementing new instructional approaches or curricula in reading/language arts/English	2.65*	1.66	0.99	1.24	0.80	0.97
Implementing new instructional approaches or curricula in mathematics	1.78	1.79*	1.68	1.39	0.56*	0.85
Providing additional instruction to low-achieving students	2.39*	3.46*	1.00	0.48*	0.31*	0.83
Restructuring the school day to teach core content areas in greater depth (e.g., establishing a literacy block)	1.85*	1.29	2.84*	1.66*	0.55*	1.18
Increasing instructional time for all students (e.g., by lengthening the school day or year, shortening recess)	1.86*	1.22	2.48*	1.77*	0.53	0.99
Providing extended-time instructional programs (e.g., before-school, after-school or weekend instructional programs)	3.54*	2.11*	2.51*	2.49*	0.46*	1.12
Implementing strategies for increasing parents' involvement in their children's education	1.86*	2.19*	2.33*	1.33	0.76	0.98
Increasing the intensity, focus, and effectiveness of professional development	1.61	1.39	2.38*	1.3	0.54*	1.00

* = Statistically significant at the 95% confidence level.

Source: GAO analysis of NLS-NCLB data.

Appendix III: List of Education Researchers

Name	Affiliation
Dr. David K. Cohen	John Dewey Collegiate Professor of Education Walter H. Annenberg Professor of Education Policy University of Michigan
Dr. Linda Darling-Hammond	Charles Ducommon Professor of Education Stanford University
Dr. Richard Elmore	Gregory R. Anrig Professor of Educational Leadership Director, Consortium for Policy Research in Education Harvard University
Dr. David Figlio	Institute for Policy Research Northwestern University National Bureau of Economic Research
Dr. William A. Firestone	Director, Center for Educational Policy Analysis; Principal Investigator, New Jersey Math Science Partnership; Professor Rutgers University
Dr. Susan Fuhrman	President, Teachers College Columbia University
Dr. Margaret Goertz	Professor Co-Director, Consortium for Policy Research in Education University of Pennsylvania
Dr. Laura Hamilton	Senior Behavioral/Social Scientist RAND
Dr. Jane Hannaway	Director of Education Policy Urban Institute
Dr. Richard Murnane	Juliana W. and William Foss Thompson Professor of Education and Society Harvard University
Dr. William Sanders	Senior Research Fellow University of North Carolina
Dr. Brian Stecher	Senior Social Scientist RAND

Source: GAO.

Appendix IV: Studies Meeting GAO's Criteria for Methodological Quality

Title	Author	Source	Method
Accountability and Teaching Practices: School Level Actions and Teacher Responses	Laura. S Hamilton; Brian M Stecher; Jennifer Linn Russell; Julie A. Marsh; Jeremy Miles	“Strong States, Weak Schools: The Benefits and Dilemmas of Centralized Accountability” <i>Research in Sociology of Education</i> , vol. 16, 2008	Case studies of three states; representative surveys for these states
Catching Up Impact of the Talent Development Ninth Grade Instructional Interventions in Reading and Mathematics in High-Poverty High Schools	Robert Belfanz; Nettie Legters; Will Jordan	Report 69 April 2004 The Johns Hopkins University Center for Research on the Education of Students Placed at Risk	Quasi-experimental design with matched groups; multiple regressions used with data. Limitations: Two school districts (around Baltimore); small percentage of all those enrolled in the 9th grade
Differentiated Curriculum Enhancement in Inclusive Middle School Science: Effects on Classroom and High-Stakes Tests	Margo A. Mastropieri; Thomas E. Scruggs; Jennifer J. Norland; Sheri Berkeley; Kimberly McDuffie; Elizabeth Halloran Tornquist; Nicole Connors	<i>The Journal of Special Education</i> vol. 40, no. 3. 2006, 130-137	Quasi-experimental design; 13 classes matched by teacher, and randomly assigned to treatment or control group. Limitations: some external validity issues
Effective Programs in Elementary Mathematics: A Best-Evidence Synthesis	Robert E. Slavin; Cynthia Lake	Review of Educational Research. Washington: September 2008. vol. 78, issue 3. 427	Literature review using a best-evidence synthesis (related to a meta-analysis)
Feeling the Florida Heat? How Low-Performing Schools Respond to Voucher and Accountability Pressure	Cecilia Elena Rouse; Jane Hannaway; Dan Goldhaber; David Figlio	Calder/Urban Institute National Center for Analysis of Longitudinal Data in Education Research Working Paper November 2007	Administrative data used to develop comparison groups of schools; regression discontinuity design; results apply to Florida schools only
Formulating Secondary-Level Reading Interventions	Debra M. Kamps; Charles R. Greenwood	<i>Journal of Learning Disabilities</i> , vol. 38, no. 6. November/December 200, 500-509	Quasi-experimental; random assignment of schools, but not students; Limitations: cannot be generalized beyond the 8 schools involved in the study
Helping At-Risk Students Meet Standards A Synthesis of Evidence-Based Classroom Practices	Zoe Barley; Patricia A. Lauer; Sheila A. Arens; Helen S. Apthorp; Kelly S. Englert; David Snow; Motoko Akiba	Regional Education Laboratory Office of Educational Research and Improvement U.S. Department of Education Mid-continent Research for Education and Learning October 2002 corrected 12/02	Literature review; in some cases a meta-analysis was conducted; effect sizes were computed for meta-analysis when available; some studies were outside the time frames of our search criteria
High Poverty Schools and the Distribution of Teachers and Principals	Charles Clotfelter; Helen F. Ladd; Jacob Vigdor; Justin Wheeler	Sanford Working Paper Series SAN06-08 December 2006	Time series analysis using administrative data for all schools in North Carolina. Limitation: applies to North Carolina only

**Appendix IV: Studies Meeting GAO's Criteria
for Methodological Quality**

Title	Author	Source	Method
High Stakes Testing and Curricular Control: A Qualitative Metasynthesis	Wayne Au	Educational Researcher; vol. 36, no. 5, June/Jul 2007; 258-267	Meta-synthesis of qualitative studies Limitations: Results for Chicago only; some coding issues
Instructional Policy and Classroom Performance: The Mathematics Reform in CA	David K. Cohen; Heather C. Hill	Teachers College Record, vol. 102, no. 2. February 2000, 294-343	Regression analysis of data from teacher surveys and administrative data. Limitations: results based on a 1994 survey; response rate was 61 percent
Instructional Time in Elementary Schools A Closer Look at Changes for Specific Subjects	Center on Education Policy	From the Capital to the Classroom: Year of the No Child Left Behind Act Center on Education Policy February 2008	Survey of school districts and states, qualitative interviews Limitation: high non-response rate from school districts in large urban areas
Standards in Classroom Practice: Research Synthesis	Helen S. Apthorp; Ceri B. Dean; Judy E. Florian; Patricia A. Lauder; Robert Reichardt; Nancy M. Sanders; Ravay Snow-Renner	Regional Education Laboratory Office of Educational Research and Improvement U.S. Department of Education Mid-continent Research for Education and Learning October 31, 2001	Literature review; no meta-analysis conducted; some studies outside our time frame
Standards-Based Reform in Practice: Evidence on State Policy and Classroom Instruction from the NAEP State Assessments	Christopher B. Swanson; David Lee Stevenson	<i>Educational Evaluation and Policy Analysis</i> , vol. 24, no. 1. Spring 2002, 1-27	Hierarchical linear modeling on survey data from the National Assessment of Educational Progress (NAEP); limitation is that only 30 of the original 40 states are included, with some of the largest of the states missing
Studying Large-Scale Reforms of Instructional Practice: An Example from Mathematics and Science	Laura S. Hamilton; Daniel F. McCaffrey; Brian Stecher; Stephen P. Klein; Abby Robyn; Delia Bugliari	<i>Educational Evaluation and Policy Analysis</i> , vol. 25, no. 1. Spring 2003, 1-29	Regression analysis; Limited to 11 sites; results small and positive, but not statistically significant
Supporting Literacy Across the Sunshine State: A Study of Florida Middle School Reading Coaches	Julie A. Marsh; Jennifer Sloan McCombs; J.R. Lockwood; Francisco Martorell; Daniel Gershwin; Scott Naftel; Vi-Nhuan Le; Molly Shea; Heather Barney; Al Crego	RAND Corporation 2008	Case study of Florida; longitudinal data analysis of data from 1997-1998 to 2006-2007 based on a survey of teachers, principals, and students in 8 middle schools
Teaching Methods for Secondary Algebra: A Meta-Analysis of Findings	Matthew Haas	<i>National Association of Secondary School Principals. NASSP Bulletin</i> , March 2005, 89, 642; Research Library 24	Meta-analysis of 35 studies

**Appendix IV: Studies Meeting GAO's Criteria
for Methodological Quality**

Title	Author	Source	Method
Test Preparation in New Jersey: inquiry-oriented and didactic responses	William A. Firestone; Lora Monfils; Roberta Y. Schorr	<i>Assessment in Education: Principles, Policy & Practice</i> , vol. 11, no.1. March 2004, 67-88	Survey, exploratory factor analysis, and hierarchical linear modeling time series; results limited to New Jersey
The Influence of Standards on K-12 Teaching and Student Learning: A Research Synthesis	Patricia A. Lauer; David Snow; Mya Martin-Glenn; Rebecca J. Van Buhler; Kristen Stoutemyer; Ravay Snow-Renner	Regional Education Laboratory, August 19, 2005	Literature review; no meta-analysis; both quantitative and qualitative studies used; comprehensive selection process
The New Accountability, Student Failure, and Teachers' Work in Urban High Schools	Dorothea Anagnostopoulous	<i>Educational Policy</i> , vol. 17, no. 3. July 2003, 291-316	Case study of two high schools; findings are suggestive
Value-Added Assessment in Practice: Lessons from Pennsylvania Value-Added Assessment System Pilot Project	Daniel F. McCaffrey; Laura S. Hamilton	RAND Corporation 2007	Quasi-experimental design for 93 non-random study districts in Pennsylvania; not generalizable to the nation or the state

Source: GAO analysis.

Appendix V: Comments from the Department of Education



UNITED STATES DEPARTMENT OF EDUCATION
OFFICE OF PLANNING, EVALUATION AND POLICY DEVELOPMENT
October 2, 2009

Ms. Cornelia M. Ashby
Director
Education, Workforce, and
Income Security Issues
U.S. Government Accountability Office
Washington, DC 20548

Dear Ms. Ashby:

Thank you for the opportunity to comment on the draft GAO report, *Student Achievement: Schools Use Multiple Strategies to Help Students Meet Academic Standards, Especially Schools with Higher Proportions of Low Income and Minority Students*.

GAO's report asks important questions about the effects of standards-based accountability on instructional practices and the effectiveness of specific instructional practices in improving student achievement, and seeks to answer these questions through a literature review and interviews with prominent education researchers. The report also examines data on the types of instructional practices that schools and teachers are using to help students achieve to state academic standards, in part based on surveys conducted for the Department's National Longitudinal Study of No Child Left Behind. While the report addresses important policy questions, there are some issues pertaining to the study's approach that we recommend be taken into consideration.

First, the draft report does not clearly explain how GAO selected the 20 studies included in its literature review or the methods used in the studies that were selected. Moreover, in discussing specific findings from the literature review, the report frequently does not indicate which studies are being relied on as evidence for each finding.

Second, the report mixes findings that may be based on rigorous research with findings that appear to be based on conjecture and on what "some researchers believe," and does not always present a complete and balanced summary of the relevant research. For example, the report states that "difficulties in aligning practice with standards were attributed, in part, to current accountability requirements," but appears to rely only on expert opinion for this causal conclusion. Similarly, the report states that "a few researchers as well as some of the literature we reviewed report some unintended negative consequences on instruction as a result of assessment practices," including the reported consequences of "multiple choice tests that do not encourage more challenging teacher practices" and "instructional practices that narrow the curriculum." These statements may accurately report the opinions of the individuals interviewed, but the report provides weak empirical evidence to support these conclusions and does not include all of the available evidence.

400 MARYLAND AVE., SW, WASHINGTON, DC 20202
www.ed.gov

Our mission is to ensure equal access to education and to promote educational excellence throughout the nation.

With respect to the assertion that the assessment provisions in the Elementary and Secondary Education Act, as amended by the No Child Left Behind Act (NCLB), have resulted in multiple choice tests that do not encourage more challenging teaching practices, the report does not present any data on the extent to which test formats have changed or on the relationship between test format and teaching practices. Instead the report notes that some researchers believe that states are increasingly using multiple-choice testing formats, and hypothesizes that teachers “may be influenced” to change their teaching approaches because of the tests. Any conclusions about what the “research shows” should be supported by specific references to rigorous research that used appropriate methods for measuring impacts.

In the discussion of whether there has been narrowing of the curriculum, the report notes that a Department survey found that 18 to 22 percent of elementary teachers reported increasing instructional time for mathematics and reading, respectively, and concludes that this is occurring “at the expense of other non-tested subjects.” However, the report does not mention the finding from the same survey that most elementary teachers reported no change from 2004-05 to 2006-07 in the amount of instructional time that they spent on other subjects. The report also notes that some research has raised concern that teachers may be restricting the breadth of content covered within a particular subject, but does not acknowledge the converse concern, based on research conducted for the Third International Math and Science Study (TIMSS), that curricula in American schools may be “a mile wide and an inch deep” and thus some refocusing of curricula may be beneficial.

The Department recognizes that improvements in assessment and accountability systems could help enable schools to strengthen instructional practices and improve student achievement. As one step toward that goal, the American Recovery and Reinvestment Act included \$250 million for Statewide Data Systems to help ensure that states and school districts have the robust data systems they need to provide information on individual student outcomes that educators and policymakers can use to drive educational improvement. More research is needed to better understand what instructional practices and policy changes could be most effective in closing achievement gaps and improving educational outcomes.

Attached are technical comments provided by Department staff on the text of the report. If you have any questions, we would be glad to discuss our comments with your research team.

Sincerely,



Alan Ginsburg
Director
Policy and Program Studies Service

Enclosure

Appendix VI: GAO Contact and Staff Acknowledgments

GAO Contact

Cornelia M. Ashby (202) 512-7215 or ashbyc@gao.gov.

Staff Acknowledgments

Janet Mascia (Assistant Director), Bryon Gordon (Assistant Director), and Andrew Nelson (Analyst-in-Charge) managed all aspects of the assignment. Linda Stokes and Caitlin Tobin made significant contributions to this report in all aspects of the work. Kate van Gelder contributed to writing this report, and Ashley McCall contributed to research for the report. Luann Moy, Justin Fisher, Cathy Hurley, Douglas Sloane, and John Smale Jr. provided key technical support, and Doreen Feldman and Sheila R. McCoy provided legal support. Mimi Nguyen developed the graphics for the report.

GAO's Mission

The Government Accountability Office, the audit, evaluation, and investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.

Obtaining Copies of GAO Reports and Testimony

The fastest and easiest way to obtain copies of GAO documents at no cost is through GAO's Web site (www.gao.gov). Each weekday afternoon, GAO posts on its Web site newly released reports, testimony, and correspondence. To have GAO e-mail you a list of newly posted products, go to www.gao.gov and select "E-mail Updates."

Order by Phone

The price of each GAO publication reflects GAO's actual cost of production and distribution and depends on the number of pages in the publication and whether the publication is printed in color or black and white. Pricing and ordering information is posted on GAO's Web site, <http://www.gao.gov/ordering.htm>.

Place orders by calling (202) 512-6000, toll free (866) 801-7077, or TDD (202) 512-2537.

Orders may be paid for using American Express, Discover Card, MasterCard, Visa, check, or money order. Call for additional information.

To Report Fraud, Waste, and Abuse in Federal Programs

Contact:

Web site: www.gao.gov/fraudnet/fraudnet.htm

E-mail: fraudnet@gao.gov

Automated answering system: (800) 424-5454 or (202) 512-7470

Congressional Relations

Ralph Dawn, Managing Director, dawnr@gao.gov, (202) 512-4400
U.S. Government Accountability Office, 441 G Street NW, Room 7125
Washington, DC 20548

Public Affairs

Chuck Young, Managing Director, youngc1@gao.gov, (202) 512-4800
U.S. Government Accountability Office, 441 G Street NW, Room 7149
Washington, DC 20548

