Chapter 2
Methodology

The classroom (grades 4 and 8) constituted the primary unit of analysis for this study. Researchers also focused some attention on the school, district, and state levels, primarily through interviews. A mix of quantitative and qualitative methods were employed. The key quantitative activity was the administration of an extensive survey about teachers' mathematics instructional practices, professional development, and professional background to 805 California teachers; researchers then statistically correlated the survey responses with the SAT-9 data of the responding teachers' students. On the qualitative side, researchers conducted classroom observations of and interviews with 55 teachers and interviewed the principals at the schools of these teachers. District- and state-level interviews were also conducted.

The study and its instruments were designed around a common core of topics based on the project's major research questions. Thus, the data yielded by the survey, interviews, and classroom observations could be triangulated to confirm and enrich the findings. Nevertheless, each of the data sources yielded some different information so as not to be completely redundant. All, however, addressed the important ideas embodied in the research questions.

Teacher Survey

A survey administered to fourth-grade teachers and eighth-grade mathematics teachers constituted one of the primary sources of data for this study. A total of 805 teachers in 11 California school districts were surveyed about their mathematics instructional practices, professional development, and professional background. The research staff sent out the questionnaires used in this survey on a rolling basis from February through May of 1999.

Selection of Districts. A purposive sample of 11 districts was selected. This sample contained six districts considered to have "large" total student enrollments, and five districts considered to have "moderate" total student enrollments. Districts were chosen to be geographically dispersed across California, and most had relatively large numbers of minority, low-income, and limited English proficient (LEP) students. Taken together, the 11 districts contained 1.2 million students—20.2% of all students in the state.

Selection of Schools. Within each of the 11 districts, a random sample of schools was selected. The number of schools selected was designed to provide a target sample of approximately 800 teachers, including (a) a higher proportion of teachers from the larger districts, since larger districts contain a higher proportion of students, and (b) more fourth-grade teachers than eighth-grade teachers, since eighth-grade teachers generally teach mathematics to multiple classes and thereby represent a greater number of students. In the largest district, the targets were 75 fourth-grade teachers and 38 eighth-grade teachers. The corresponding targets in the other five large districts were 50 and 25, and in the moderate-size districts they were 40 and 20.

A systematic sampling procedure was used to select, within each district, a diverse set of schools in terms of student socioeconomic status, ethnicity, and language proficiency. Schools with fewer than 10 fourth- or eighth-grade students were excluded, as were alternative and community schools. Elementary schools selected for the Evaluation of California's Class Size Reduction Program also were excluded, so as to avoid an excessive burden on teachers. The number of schools selected as candidates for participation totaled 168 elementary schools and 79 middle schools.

Once schools were selected, research staff contacted the principals of the selected schools to obtain their agreement to participate in the study. Several of the initially selected schools, however, declined to participate and, as possible, were replaced with other schools of similar demographic profile. The total number of schools ultimately included in the sample was 158 elementary schools and 68 middle schools.[1]

Selection of Teachers. Within each school in the sample, questionnaires were sent to all of the fourth-grade teachers and all of the eighth-grade mathematics teachers. (Teacher names were obtained from the school principal, and the questionnaires were mailed directly to each teacher.) In sum, questionnaires were sent to 570 fourth-grade teachers and 235 eighth-grade teachers.

Questionnaire Development. The questionnaire was based on other, pre-existing survey instruments of similar nature, namely: (1) the "Survey of Elementary Mathematics Education in California" questionnaire developed by the Center for Research on the Context of Teaching at Stanford University; (2) questionnaires developed by Horizon Research, Inc. for the National Science Foundation's Local Systemic Change Initiative, and (3) the "Reform Up Close" questionnaire developed by the Wisconsin Center for Education Research. Once drafted, the questionnaire underwent numerous rounds of revision based on feedback from project staff, Advisory Group members, and CDE staff.

Two different versions of the questionnaire were developed, one for the fourth-grade teachers and one for the eighth-grade teachers. Most items on the two versions were

[1] More detailed information on the school sampling procedure is included in the RAND report in Appendix A.

identical; however, there were some differences necessary given that while most fourth-grade teachers teach mathematics to only one group of students, eighth-grade teachers often teach multiple mathematics classes per day. Because any given teacher may use different practices in different classes taught, the eighth-grade version instructed respondents to fill out the practices questions for only one class: their "first mathematics class of the day in which at least half of the students are in 8th grade." Teachers were then asked to indicate the class period for which they were filling out the questionnaire, and to write in the title of this class (e.g., Math 8, Algebra, Integrated Math, etc.).

Questionnaire Composition. The questionnaire was mainly composed of discrete-answer questions with a few open-ended response items. The items on the questionnaire were divided into the following topic areas:

- Current teaching situation: grade levels taught, number of classes per day taught, and subjects other than mathematics taught
- Mathematics instruction "in your class" (fourth-grade) / "in a particular class" (eighth-grade): amount of time for mathematics instruction, class size and class composition, frequency of use of a wide range of instructional practices (on a 5-point Likert scale, from "never" to "almost daily"), objectives for mathematics instruction, mathematics content topics taught, and curriculum materials
- Recent developments in mathematics education: familiarity with various standards documents, opinions about these documents, and ratings of school / district alignment with the documents (on a 4-point Likert scale, from "disagree strongly to "agree strongly," with a fifth option for "don't know")
- Professional development and support: amount of mathematics professional development (total and by certain topics) since January 1998, opinions about support, and frequency of teacher collaboration
- Professional background: mathematics courses taken, degree received, teaching credential, and years of teaching experience
- Teacher demographic information: gender and racial / ethnic background
- Additional comments: open-ended items about factors facilitating or impeding effective mathematics instruction.

The complete questionnaire (both fourth-grade and eighth-grade versions) is included in Appendix B.

Response Rate. Questionnaires were received back from 310 (54.4%) fourth-grade teachers and 139 (59.1%) eighth-grade teachers. However, 49 of these questionnaires were eliminated due to the following reasons:

- the respondent's class did not contain at least one-third students at the appropriate grade-level (fourth or eighth)
- the respondent had not been teaching for most of the school year

- the students of the respondent were lacking test scores
- at the fourth-grade level, the respondent was part of a team where different teachers shared or rotated students for mathematics instruction (meaning that students' test scores could not be linked to a particular teacher's instruction)
- the students in the respondent's classes could not be identified by project staff.

After these eliminations were made, questionnaires remained from 281 (49.3%) fourth-grade teachers from 136 schools and 119 (50.6%) eighth-grade teachers from 57 schools.[2]

Generalizability. Because the participating districts were not a random sample of all districts in California and because of the moderate response rate on the survey, the results of this study may not be representative of all the state's students and teachers. This is especially true for districts with small enrollments. Consequently, the relationships (or lack thereof) presented in this report cannot be generalized to the state as a whole. Nevertheless, due to the large number of students and teachers included in the sample, the results are likely to be meaningful and merit further consideration.

Student Achievement Data

The research design called for the linking of teachers' questionnaire responses with mathematics achievement data of their students to see if any correlations between practices and achievement existed. The student mathematics achievement data selected for use in this analysis were from the Stanford Achievement Test, Version 9 (SAT-9), a multiple-choice assessment administered to nearly all California students in grades 2–11. Students took this test in the spring of 1999, after they had been in the class of the participating teacher for most of the year.

Participating districts provided the data. Some districts were able to provide the student data given only teachers' names. Other districts required student identification numbers; in these districts, researchers obtained the class rosters of the teachers who had responded to the survey. A small number of rosters could not be obtained, so the questionnaires for these teachers had to be eliminated from the study.

The 281 fourth-grade teachers had a total of 6,885 students with valid SAT-9 scores. However, 70 of these students were missing demographic data and were excluded from further analyses, so the final fourth-grade sample consisted of 6,815 students. The 118 eighth-grade teachers included in the survey-test score linking analysis had 3,063 students,

[2] One of the 119 eighth-grade teachers filled out the questionnaire about a geometry class. Because this was the only geometry class in the sample, it was excluded from the analysis linking practices with test scores. However, this teacher was kept in the sample for most other analyses.

but 30 were missing demographic data, resulting in a final eighth-grade sample of 3,033 students.

The student demographic data included in the analysis consisted of gender, racial/ethnic group, home language, and whether the student participated in a gifted program, a special education program, and/or a free or reduced price lunch program. Students' 1998 SAT-9 mathematics scores and their 1998 and 1999 SAT-9 reading scores were included in the analysis as well. (See the RAND report in Appendix A for a description of how these data were used.)

School Visits: Classroom Observations and Interviews

In May and June of 1999, trained mathematics observers visited the classrooms of and conducted interviews with 55 teachers in the study. All of the teachers had filled out the questionnaire and were located in eight of the eleven districts participating in the study.

Selection of schools/classrooms for visits. Eight of the eleven study districts were selected for school visits. Within each district, the goal was to select two elementary schools and two middle schools to visit, and to observe and interview two teachers in each selected school, thereby yielding a sample of 64 classrooms observed. The procedure for selecting the visited schools/teachers was as follows:

- Within each district, all schools from which at least two teachers had returned the questionnaire were identified.
- If there were more than two such schools in the district, researchers randomly selected two from the list.
- The questionnaires from the teachers at the selected schools were screened (a) to make sure their classes consisted of at least half fourth or eighth graders and (b) to make sure that the observation sample as a whole would include a wide range of class types (e.g., at the eighth-grade level, not too many algebra classes; at both grade levels, not too many high-percentage LEP classes).
- For any school that did not have at least two teachers' classes meet the selection criteria, researchers randomly selected a replacement school and screened it similarly.
- The selected schools/teachers were contacted to request the visit. Schools that declined were replaced with others, using the same random selection and screening criteria. Teachers were offered a $25 honorarium for participation.
- For schools from which more than two teachers returned the questionnaire, two of the teachers were selected based on convenience factors (or, if possible, more than two teachers were visited/observed). At the eighth-grade level, efforts were made to

visit the exact class periods about which the teachers filled out the questionnaire, or, if this was not possible, to visit a "similar" class.

Fifty-five teachers—28 fourth-grade teachers from 14 elementary schools and 27 eighth-grade teachers from 14 middle schools—were visited and interviewed. The principals at 26 of the 28 schools also were interviewed.

The visited schools displayed a wide range of demographic characteristics and overall student achievement. For example, several different Academic Performance Index (API) rankings—both statewide rank and similar schools rank—were represented among the visited schools. Figure 2.1 shows the API rankings of the visited fourth-grade schools, and Figure 2.2 shows the API rankings of the visited eighth-grade schools.

Figure 2.1
Fourth-Grade Visited Schools' 1999 Academic Performance Index Rankings

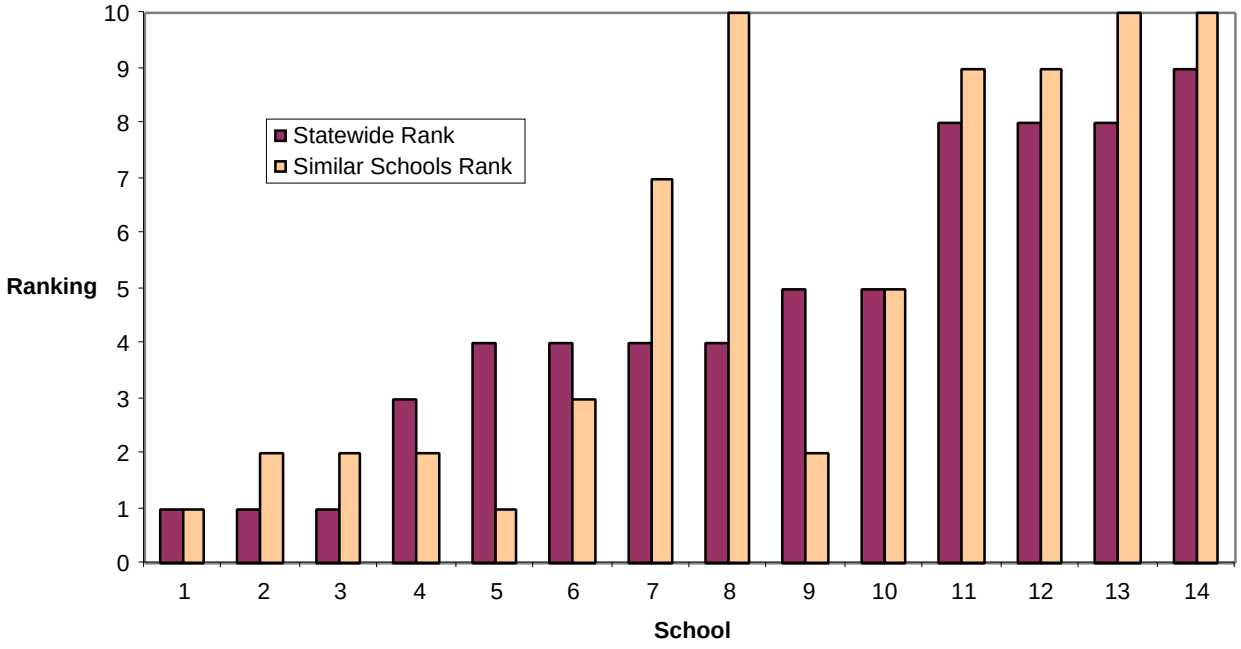
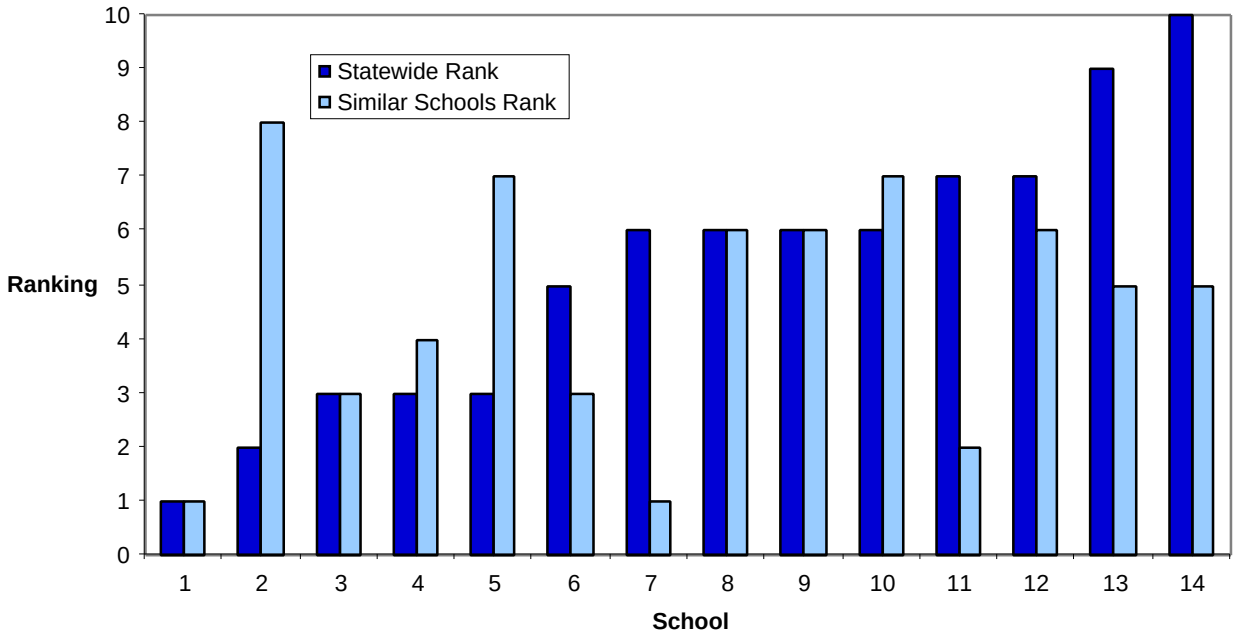


Figure 2.2
Eighth-Grade Visited Schools' 1999 Academic Performance Index Rankings



Classroom observations. Each teacher was visited only once, and only one mathematics lesson / class was observed.[3] In most cases the visit was made by a single observer, but in some cases two observers made the visit. Observers were asked to write up a summary of the observation, including attention to the content of the lesson; the organization of students; the purpose of the lesson; representations, tools, and resources used; assessment during the lesson; focus of classroom discourse; language differences; students with special needs; and behavior and discipline. The complete protocol for this write-up is included in Appendix B.

Observers were also asked to compare each teacher's practice *as observed* to practice as reported by the teacher on the questionnaire. The purpose of this comparison was to validate the questionnaire. However, because most of the questionnaire items about teaching practices asked about frequency of their use, complete validation was not possible given the "one-shot" observation. Observers could, nevertheless, attempt to verify the presence of practices teachers reported engaging in "almost daily," and, conversely, verify the absence of practices teachers reported "never" using. The overall results of this analysis, across all the classroom observations, did not find the questionnaire to be invalid. (Two

[3] That each class was observed only once is a limitation of the study, as instruction in that one class may not have been representative of the teacher's instruction. That the visits were made toward the end of the year compounds this problem, as instruction close to the end of the year may differ from instruction earlier in the year.

questionnaire items were, however, found to be ambiguous and were thus not included in the analysis.)

Teacher interviews. The teachers' whose classes were observed were also interviewed. A brief interview was conducted prior to the observation and consisted mainly of questions about the lesson planned. A lengthier interview was conducted following the observation and included questions about the lesson observed, the teacher's "philosophy and practice" regarding mathematics instruction, perceived influences on mathematics instruction, and effectiveness in teaching mathematics. Again, the complete protocol is included in Appendix B.

Principal interviews. As mentioned, the principals at most of the observed schools also were interviewed. The principal interview protocol, also included in Appendix B, contained questions about the school's mathematics program, support from the district, school and teacher discretion, influences on mathematics instruction and achievement, professional development, and areas for improvement.

District-Level Interviews

In four of the eight districts where school visits occurred, a district curriculum administrator (e.g., district mathematics coordinator) was interviewed. The district-level interview included questions about the district's mathematics program; influences on mathematics instruction in the district; the use of content standards; professional development; student mathematics achievement; strengths and weaknesses of district mathematics instruction; and accountability. The district-level interview protocol is included in Appendix B.

Other Interviews

Interviews were conducted with a variety of other stakeholders as well, in order to gain a wide range of additional perspectives on mathematics instruction and implications for policy. Individuals who were interviewed included members of the Legislature / legislative staff, members of the State Board of Education and their staff, administrators from the California Mathematics Project and the California School Board Association, a mathematics professor, and a focus group of teachers formed by the California Federation of Teachers.

Questions in these interviews solicited opinions on the current level of mathematics achievement in California, on the appropriate role of state policy makers for the

improvement of mathematics instruction, and on the appropriateness of the current state strategy for improving mathematics achievement. Interviewers next presented some of the study's major findings and asked for opinions on the appropriate policy responses to these findings. Because the interviews involved discussion of study findings, they took place toward the end of the study, in April and May of 2000.

Tenth-Grade Pilot Study

In addition to the more thorough investigation at grades 4 and 8, some exploratory research and development work was conducted at grade 10. This exploratory work utilized instruments and protocols employed at the fourth- and eighth-grade levels, but did not analyze any student achievement data. The intent of this work was to refine the instruments and procedures for use in a future high school study and to frame the major issues involved in undertaking such a study. The tenth-grade research consisted of the following elements:

- The eighth-grade teacher questionnaire was adapted for the tenth-grade level. (See Appendix B.) Four mathematics teachers from two high schools within a single district completed the instrument. These teachers then participated in a focus group to critique the questionnaire and its appropriateness for use with high school teachers.

- Researchers conducted observations in classrooms of the four teachers who had participated in the focus group. The classes that were observed—two Geometry classes, one Algebra 1 class, and one Advanced Algebra class—each had more than 50% tenth-grade student enrollment. The observation protocol was the same as that used for the fourth- and eighth-grade levels.

- Interviews were then conducted with the four teachers, the mathematics department chair and principal at each school, and the district mathematics resource specialist. The protocols for these interviews were similar to those used for the larger study.

The findings from this exploratory study are not included with those from the main study in the body of this report. Rather, the implications for a tenth-grade study are included in Appendix D.

(This page intentionally left blank.)