New Tools for Analyzing Teaching, Curriculum and Standards in Mathematics & Science

Results from Survey of Enacted Curriculum Project
Final Report

July 2001

Rolf K. Blank
Andrew Porter
John Smithson

Project Team

- Council of Chief State School Officers, State Education Assessment Center
- Wisconsin Center for Education Research at the University of Wisconsin, Madison
- State Collaborative: Iowa, Kentucky, Louisiana, Massachusetts, Minnesota, Missouri, North Carolina, Ohio, Pennsylvania, South Carolina, West Virginia

Project funding provided by a grant from the National Science Foundation (REC98-03080), and support from the collaborative states.

COUNCIL OF CHIEF STATE SCHOOL OFFICERS
ONE MASSACHUSETTS AVE., NW, WASHINGTON, DC 20001-1431 TEL.: (202) 408-5505

The Council of Chief State School Officers (CCSSO) is a nationwide, nonprofit organization of the public officials who head departments of elementary and secondary education in the states, the District of Columbia, the Department of Defense Education Activity, and five extra-state jurisdictions. CCSSO seeks its members' consensus on major education issues and expresses their view to civic and professional organizations, to federal agencies, to Congress, and to the public. Through its structure of standing and special committees, the Council responds to a broad range of concerns about education and provides leadership on major education issues.

The State Education Assessment Center was established by chief state school officers to improve the information base on education in the United States, especially from a state perspective. The Center works to improve the breadth, quality, and comparability of data on education, including state-by-state achievement data, instructional data, indicators of education quality, and performance assessment of teachers and students.

ISBN: 1-884037-73-9
Copies of this report may be ordered for $15.00 per copy from:
Council of Chief State School Officers
Attn: Publications
One Massachusetts Avenue NW Suite 700
Washington, DC 20001-1431
Phone: (202) 336-7016
Fax: (202) 408-8072
or, go to www.ccsso.org\Publications

About the Project Team

This report summarizes the results from a two-year collaborative project that developed and tested the use of Surveys of Enacted Curriculum for analyzing mathematics and science education in the classroom. The project was carried out by the work of many people and several organizations:

- **CCSSO:** As the lead organization, the Council of Chief State School Officers formed the study team, gained cooperation of states and support of the National Science Foundation, and managed the components of the study. CCSSO is responsible for the study products and future use and application of the materials developed in the project.
  Rolf K. Blank, Principal Investigator/Project Director; Linda Dager Wilson, Jennifer Manise, Cynthia Dardine, Barbara Brathwaite, Doreen Langesen.

- **WCER:** Wisconsin Center for Education Research led the design and development of the survey instruments and the alignment study. The Wisconsin Center also managed data input, editing, and analysis, and designed the reporting formats and use of software.
  Andrew Porter, WCER Director and Senior Researcher; John Smithson, Project Manager; Molly Gordon, Eleanor Cameron, Alissa Minor.

- **State Collaborative:** State education representatives from the 11 state partners worked on selecting schools and districts for study participation, provided followup for data collection, assisted in survey design and analysis, and recommended approaches to reporting study results.
  State Staff: *Iowa*: Tony Heiting, Matthew Ludwig; *Kentucky*: Ann Bartosh, Carol Hanley; *Louisiana*: Jennifer Falls, Gary Kratzer; *Massachusetts*: Anne Collins, Joyce Croce, Beth McBride, David Perda; *Minnesota*: Ed Hessler, Sharon Stenglein; *Missouri*: Greg Steuwe-Portnoff, Walt Brown; *North Carolina*: Mike Kestner, Linda Patch; *Ohio*: Anne Mikesell, Stan Santilli; *Pennsylvania*: Robert Livingston; *South Carolina*: Lane Peeler, Linda Sinclair; *West Virginia*: Phyllis Barnhart, Larry Lamb, Brenda West.

- **NSF:** The survey instrument development, study design, and collection and analysis were supported by a grant from the National Science Foundation, Education and Human Resources Directorate, Division of Research, Evaluation and Communication (REC 98-03080). We appreciate the advice and encouragement of NSF/REC Program Officers, Bernice Anderson and Larry Suter.

Table of Contents

[Available on compact disc from CCSSO, or on web: www.ccsso.org]:

Survey of Classroom Practices (Survey Instruments): Mathematics, Science (Elementary, Middle Grades, High School)

Complete Data Charts for Reporting 1999 Survey Data: Mathematics and Science

List of Tables and Charts

Executive Summary

The Council of Chief State School Officers (CCSSO) has completed a two-year project to develop and test ***Surveys of Enacted Curriculum in Mathematics and Science***. This final project report describes advances in survey methods and analyses of enacted curriculum data; it highlights central findings of the research; and it presents important applications of the survey and data tools for linking education research and improvement of practice. The project was supported by a grant from the National Science Foundation (NSF) (REC 98-03080). The study was led by CCSSO staff and researchers at the Wisconsin Center for Education Research (WCER). Schools, teachers, and state specialists from 11 states participated in the study.

Schools across the nation are working to adapt and improve curricula and teaching practices to meet the standards for learning established by states and school districts. In mathematics and science education, "standards-based reform" typically means that teachers must plan and implement their curriculum and teaching in relation to state or district content standards by subject. Standards often include specific, challenging expectations for student knowledge and skills. A major question for education decision-makers is how best to assist teachers in improving their curriculum content and teaching practices, with the ultimate goal of improving student achievement. An important question for researchers is how best to measure change in instruction, related to standards, and determine the relationship of changes in teaching to student learning.

The Surveys of Enacted Curriculum project was designed to address these broad questions about standards-based reform by testing a survey approach to analyzing the enacted curriculum in math and science. We defined "enacted curriculum" as the actual subject content and instructional practices experienced by students in classrooms. Four primary results from the study are highlighted:

1. **Demonstrated efficient, reliable method of data collection**. The Surveys of Enacted Curriculum provided educators and researchers with reliable, comparable data on the curriculum that is taught in math and science classrooms. The teacher self-report survey design proved to be effective for collection of curriculum data at all grade levels. We identified methods to raise survey response rates, which had lowered as a result of the length and complexity of the survey.

2. **Advanced survey methodology on curriculum and instruction**. Our methods of surveying and analyzing curriculum data include several critical advances from prior survey approaches. First, data

on curriculum content were collected with a two-dimensional matrix-- the intersection of content topic (e.g., fractions) and expectations for learning (e.g., solve non-routine problems). Second, the surveys were designed with sufficient depth and specificity of questions about classroom practices to capture a high degree of variation among classrooms and schools. Third, the surveys are comprehensive so that pedagogy and curriculum can be analyzed by teacher preparation, quality of professional development, and school climate and conditions for teaching.

3. **Produced findings on state initiatives and developed tools for data reporting**. The study results indicate that the curriculum taught in mathematics and science differed according to the amount of teacher preparation in math and science through professional development and according to levels of implementation of state math and science reform initiatives. Unique methods of aggregating and reporting survey data were developed in the project, which allowed in-depth analysis of differences in subject content and pedagogy across schools, districts, and states. The data report formats developed in the project were aimed toward teachers and administrators.

   We developed and used a topographical mapping software for reporting on math and science content that pictorially displays central differences in content taught over an academic year. Bar graph formats for data reporting were also designed with input from state specialists. Item profiles and summary scales were developed for accessibility in analyzing differences in pedagogy, use of materials, school conditions, and teacher perceptions.

4. **Conducted alignment analyses with important applications for teachers and policy-makers.** The survey data from teachers on their enacted curriculum were compared to the content of state student assessments in math and science. The content matrix (topics by expectations) proved to be effective in both collecting data from teachers and for systematically categorizing the test items found on state assessments. The procedures and tools for surveying teachers and analyzing content of assessments provide very strong potential for application in districts and states that want to determine the progress of standards-based math and science improvements. However, it is critical to observe that alignment analyses possible through these tools are as important for teachers themselves to reflect and improve their own practice, as they are for policymakers and leaders to determine the extent of change in practice across schools and classrooms.

Study Products

In addition to this report, the following study products were completed by CCSSO and the project partners:

- *Survey of Classroom Practices: Mathematics, Science* (Paper and electronic forms available): For elementary, middle, and high school levels. Survey instruments for teachers focused on instructional practices, subject content, teacher preparation/professional development, and school conditions (1999).

- *Using Data on Enacted Curriculum in Mathematics and Science*: Summary Report—initial report of findings; 48 pages, 15 full-page data charts. CCSSO (2000).
  Also, State Reports with complete data on samples from each state: 50 full-page data charts, 11 states. Available by request.

- *A Guide for Professional Development*: Designs and materials for five Professional Development Workshops on Use of Surveys of Enacted Curriculum for educators and administrators; 50 pages. CCSSO (2001).

- *Surveys of Enacted Curriculum* (compact disc): Electronic versions of all survey instruments, reports, appendices, background papers, data analysis programs, report formats, and study products. CCSSO/WCER (2001).

Introduction

In Spring 1999, schools and teachers in 11 states participated in a Study of the Enacted Curriculum in Mathematics and Science classrooms. More than 600 teachers across the states completed self-report surveys that covered the subject content they taught and the instructional practices they used in their classes. The goals of the study were to:

- Measure differences in instructional practices and curriculum content across a large, multi-state sample of schools and teachers;
- Determine if there are consistent differences in mathematics and science instruction that are related to state policy initiatives and state standards;
- Demonstrate the use of "surveys of enacted curriculum" to analyze classroom practices and to produce useful analyses and reports for educators.

The Study of Enacted Curriculum was a collaborative effort involving staff of CCSSO's State Education Assessment Center, researchers from the University of Wisconsin-Madison, and state education specialists in science, mathematics, and assessment. Key steps in the study included development of valid survey instruments for measuring instructional practices in math and science classrooms, collection and analysis of data, and design reporting formats and summary scales that communicate key findings to educators. The project received grant support from the National Science Foundation (NSF).

This final report describes the overall results of the study and outlines how Surveys of Enacted Curriculum (SEC) can be used for analyzing the implementation of systemic, standards-based reform in mathematics and science education. The report demonstrates how the survey and data tools can be used and explains some of the findings from the 11-state study, and it identifies how the enacted curriculum data can be used by policy-makers, administrators, resource people, teachers, and the general public. Finally, this report explains the procedures for use of the SEC, including administration and collection of the enacted curriculum data, summary scales and other measures for reporting, and analytic strategies that can be employed using the enacted curriculum data to analyze and evaluate reform initiatives.

The types of summary data and charts displayed in this report can be produced for an educational system that decides to conduct the SEC with all teachers at given grade levels or with a randomly- selected, representative sample of teachers. The kinds of results reported for the 1999 sample of teachers and schools illustrate the potential for the SEC in future applications by educational systems.

The report is organized in four chapters with three appendices. Chapter 1 examines some of the issues that arise when attributing student outcomes to policy initiatives. In light of these issues, the chapter provides a theoretical framework and rationale for use of the SEC, particularly as a methodology for analyzing effects of standards-based education reforms in math and science education.

Chapter 2 provides an overview of the survey instrument design and data analysis plan with several practical examples of how the SEC measures can be used to produce descriptive, empirical evidence for evaluating the effects of policy initiatives and professional development on instructional practice.

Chapter 3 demonstrates how the SEC measures can be used at the local level to assist efforts to improve instructional practices. Using data from two schools participating in the 11–state study, the chapter presents examples of how the SEC measures are analyzed and displayed at the school and teacher level, with discussion of the kinds of analytic strategies that can be used in a district or school setting by local specialists and professional development leaders.

Chapter 4 concludes the report with discussion of issues related to the quality of SEC data. In addition to specific characteristics of the 11–state study data set, the chapter provides more general discussion regarding the validity and reliability of survey instrumentation and teacher self-report data. Suggestions are offered for insuring the quality of data in administration of the SEC by schools, districts or states.

Appendix A presents a table of descriptive data for the 1999 sample of teachers and schools. Appendix B lists the state initiatives in science and math used as the basis for sample selection. Appendix C: Analysis Guide includes sample SEC sections, Content Map interpretation, and items comprising scales for math and science. Math and science teacher surveys for elementary, middle, and high school grades are available on a compact disc, with an electronic version of this report. The CD includes a complete set of math and science charts for reporting SEC data for the teachers and schools in the 1999 study, which illustrate the range of information and applications of the Survey methodology.

Chapter 1: Analytic Challenges to Studying Systemic, Standards-Based Reform

Educational standards, accountability, and systemic approaches to reform have emerged in the past decade as favored policy tools for promoting "world-class" public education for all students. State and federal government have invested large amounts of money in developing standards, high-stake assessments, professional development, and other capacity building resources and expertise in order to improve the quality of education. Not surprisingly, policy-makers have become very interested in evaluating the effectiveness of the various funded strategies for improving educational quality. The criteria for success in such evaluations is typically outcome measures, such as student achievement scores or reduction in the achievement gap between student race/ethnic groups. However, it is critical to measure and analyze the quality of classroom practices, which must change if student achievement is to improve.

Complex Interactive System

While outcome measures are valuable indicators for the health of the educational system, attributing those measures to some particular policy initiative or pedagogical approach is no small task. Not only is the educational system a complex organizational entity, it is also a system dependent upon the interactions and relations of human agents at every level. While the standards-based systemic approach to reform has provided a set of tools for bringing elements of this complex system into better alignment toward a common goal, the system remains extremely complex. Investigating the effects of practice and policy requires considerable care, expertise, and investment in research and analysis (see following examples of research and evaluation with systemic reform: Zucker, et al., 1998; Clune, 1998; Corcoran, et al., 1998; Kahle, 1999; Massell, 1997; Webb, 1999; Klein, 2000; CCSSO, 2000; Systemic Research, 2000).

Simplifying the Causal Chain. K-12 education presents an exceptionally complex system with numerous steps in the causal chain between goals and initiatives for reform and student achievement. One way to simplify the causal chain is to divide the system into three components: the intended curriculum, the enacted curriculum, and the learned curriculum (i.e., student outcomes). The logic behind this chain of causality suggests that the intended curriculum, represented by policy tools consisting of content standards, curriculum frameworks, guidelines and state assessments, has effects on instructional practices and curriculum content in classrooms (enacted curriculum), which in turn impacts student learning. The concept of analyzing the types or levels of curriculum has been used consistently in the studies conducted by the International Association for Evaluation of Education (IEA), such as the recent science and mathematics studies (Beaton, 1996 a, b; Schmidt 1996 a, b; Martin, 2001; Mullis, 2001).

The data needed to measure change in student outcomes due to policy initiatives are: a) evidence that policies have changed practice in the ways predicted; and b) evidence of a relationship between change in teaching practice and student achievement. Both kinds of data are necessary in order to draw the link between policy initiative and student achievement.

In this project, we have been able to show that the Survey of Enacted Curriculum and related data analyses provide the necessary sets of data to trace a causal chain for K-12 education from policy initiatives to achievement.

Evaluating Pedagogy in relation to Subject Content. Because there is a large number of potential measures of instructional practice in schools, criteria are needed to decide what to measure. One key criterion for priorities for analyzing instruction should be utility for predicting student achievement. Two areas of instructional measures are pedagogy used in classrooms and content taught.

A focus of interest among researchers and educators is measuring the effects of the different pedagogical approaches used for instruction. For example, teacher surveys with the NAEP mathematics and science assessments have reported on the use of lecture, small group work, hands-on activities, and other teaching practices. The NELS88 study included questions for teachers aimed at measuring differences in teaching practices in math, science, and other subjects. These studies are useful in providing indicators of practices across the nation; however, the data does not indicate that the use of one or more specific teaching practices produces, by itself, improved student achievement. There are a number of measurement issues with analyzing teaching practices, including the degree of specificity of the questions for teachers (so that items accurately differentiate between teaching practices), measures of the quality of delivery of pedagogy, and the prior preparation of the teacher. Even if these measurement problems were resolved, as an independent measure, it is difficult to identify the superiority of one pedagogical approach used by a teacher as compared to other approaches (Westat/Policy Studies Associates, 2000).

One reason the connection between pedagogy and achievement has been weak is that studies have not had the tools available to control for a critical element in analyzing classroom differences -- the content of instruction. If researchers can reliably collect data across schools and classrooms on the content being taught, the relative merits of various pedagogical practices will become amenable to analysis and evaluation. The recent Third International Mathematics and Science Study (TIMSS) demonstrated the value of research tools that allow for 1) the analysis of differences in curriculum content and pedagogy, and 2) the relation of instruction to standard measures of student achievement (NCES, 1996, 1997, 1998; Beaton, 1996; Martin, et al., 2001a, b).

Potential Solutions: Tools and Strategies

The need of state and federal agencies for investigation and evaluation of reform efforts has led researchers to develop promising tools and analytic strategies for investigating systemic reform. The present study builds on prior research and development work by CCSSO and WCER including the Reform Up Close study (Porter, al., 1993), and a five-state field study of science teaching practices (CCSSO, 1997), both supported by NSF. Survey items for the present study on teaching practices and teacher preparation were previously field tested and used by Horizon Research in national surveys and evaluations of Local Systemic Collaboratives, also supported by NSF (Weiss, 1994; see www.horizon-research.com for surveys).

The content matrix design and classroom practices items of the Survey of Enacted Curriculum developed by researchers at CCSSO and WCER for the present study were also used in the 1999 National Evaluation of the Eisenhower Math/Science Program (Garet, et al., 1999; Porter, et al., 2000), and they are presently being used in the design for the National Study of Title I Schools conducted by RAND for the U.S. Department of Education. Additionally, the SEC is being used in separate evaluation studies of urban systemic initiatives by Systemic Research Inc. (2000) and the University of South Florida (USF). Together these studies are beginning to provide cumulative evidence on instructional practice and content being taught in schools around the country. By doing so, they provide researchers and policy makers valuable information that can be used within multiple analytic frameworks.

Purpose of the SEC Instruments. The survey approach used in this study offered a practical research tool for collecting consistent data on mathematics and science teaching practices and curriculum based on teacher reports of what was taught in classrooms. The enacted curriculum data give states, districts, and schools an objective method of analyzing current classroom practices in relation to content standards and the goals of systemic initiatives. The methods of aggregating and reporting survey data allow educators to analyze differences in classroom practices and curriculum among schools with varying characteristics. Districts and states can analyze differences in classroom curriculum related to state policy initiatives, state or district standards, or assessments in math and science.

The data were collected using written surveys that relied on teachers to self-report, and they were designed for elementary, middle, or high school teachers. Teachers were asked to report on the range of practices and subject areas covered during the course of the school year and to provide information on the school, class and their own professional development and preparation for teaching.

Issues Addressed. The major concepts underlying the SEC design were drawn from state and national content standards, state initiatives in science and mathematics education, and prior research studies on classroom instructional practices and curriculum content. The SEC is intended to answer many of the key questions educators and policy-makers have about patterns and differences in classroom curriculum and instructional practices across classrooms, schools, districts, and states. The following listing of major concepts from the SEC reflect the types of issues and questions that can be explored using the enacted curricular data:

- Active Learning in Science
- Problem Solving in Mathematics
- Mathematics and Science Content in Classrooms (reported by grade level)
- Multiple Assessment Strategies in Math and Science
- Use of Education Technology and Equipment
- Teacher Preparation in Subject
- Quality of Professional Development
- Influences of Policies and Standards on Practice
- Alignment of Content Taught with State Assessments

Linking the Intended and Enacted Curriculum. An important role of data on enacted curriculum is to analyze the curriculum that is taught in relation to the intended curriculum, which is typically established by state or district policies. If comparable, quantifiable descriptions are available, the research analysis can generate a measure of agreement, or alignment, between the intended and enacted curricula. The degree of correlation between the intended curriculum and the enacted curriculum in classrooms provides decision-makers and teachers with an indicator of how well the teaching is reflecting the system goals set for learning. Further, if data on enacted curriculum were collected at multiple points in time for the same teachers, the surveys could be an even stronger tool in determining the degree to which teaching practices and content are moving in the direction envisioned by the policy goals. If longitudinal data are collected for teachers and schools, analyses can determine the direction of change over time towards greater alignment with the intended curriculum.

Porter (1998) described a model for predicting the effects of education policy on change in instruction. In this model, policy tools are described on the basis of four characteristics: prescriptiveness, consistency, authority, and power. *Prescriptiveness* indicates the extent to which policy instruments, such as standards or curriculum guides, specify desired practice. *Consistency* describes the extent to which policy instruments are mutually reinforcing (i.e., aligned). For the purposes of this discussion, one important measure of consistency is the extent to which the content standards and statewide student assessments of a given state present consistent educational goals for instruction. *Authority* refers to the extent to which policies are persuasive in convincing teachers that their intent is appropriate. A curriculum policy instrument has *power* to the extent that rewards and sanctions are tied to compliance with the policy. High stakes tests are one notable example of a curricular policy with power.

The model presents a theory that can be tested, i.e., the more curriculum policies reflect these four characteristics, the stronger the influence policies will have on instructional practice. Thus, if:

1. a specific policy or set of policies are shown to be strong on three or four of the policy characteristics, and
2. data about instruction reveal substantial agreement between the intended and enacted curricula, and
3. this level of agreement has increased over the time period in which policies have operated in schools; then,
4. the evidence is supporting the premise that policies did produce change in instruction.

Analyzing Student Achievement. In order to further extend the causal model of change in education systems to include improvement of student achievement, evidence is needed to make the link between instructional practice and gains in student learning. While achievement scores alone provide some measure of the level of knowledge students have attained, the scores do not necessarily indicate when and how the knowledge was acquired. In order to measure the contribution of instructional practice to scores, a more narrow measure of achievement is necessary. By focusing on *gains in student achievement*, rather than simply the raw scores on a test at a given time, it is possible to examine the contribution of classroom experience to student achievement over specified periods of time. Measuring change in classrooms over time is necessary to demonstrate the effects of recent changes in policy and instruction on achievement.

In addition to controlling for prior achievement, such as using learning gains, the analysis must include a control variable for student socio-economic status (SES). In a recent study of the Prospects program, Rowan (1999) found that prior achievement and SES accounted for as much as 80 percent of the variance in mean achievement among classrooms. Rowan estimates the percentage of variance among classrooms to be 11 percent after controlling for prior achievement and SES. This suggests that the extent to which the classroom experience of students in a given year contributes to their overall achievement score is relatively small compared to prior achievement and SES. However, Rowan also notes that the percentage of variance attributable to classroom differences may be significantly higher when a measure of the degree of alignment between the test being given and the classroom instruction is taken into account.

With comparable data on the content of instruction and the content of an assessment instrument, an alignment variable can be calculated. If alignment succeeds in predicting student achievement above and beyond the control variables, then an argument can be presented that variance in instructional practice does cause gains in student achievement.

Chapter 2: Evidence for Systemic Reform

In this chapter, we present an overview of the study design and results of several analyses with enacted curriculum data that were collected in 1999. The results reported do not purport to be generalizable beyond the sample of teachers and schools in the 11 states that volunteered for the study. Moreover, it is important for the reader to note the sample size for the various charts employed in reporting the data results. The results are intended to be illustrative of the types of information and analyses that can result from utilization of the SEC instruments and analysis procedures. Chapter 4 examines issues related to the quality of the data, both in general terms and with respect to the data set upon which the results reported here are based.

Overview of Study Design

The survey design and instruments for the SEC in mathematics and science conducted in 1999 were based on earlier research, development, and field testing carried out collaboratively by CCSSO and WCER staff (CCSSO, 1998; Martin, et al., 1996; Smithson, Blank, & Porter, 1995). In addition, CCSSO had worked with state education leaders in developing content standards and assessments in science and mathematics (Blank, et al., 1997). WCER researchers had tested the validity and usefulness of a survey approach to collecting reliable, comparable data on classroom curriculum and practices (Smithson and Porter, 1994).

The movement of states toward standards-based reform in mathematics and science produced strong interest in reliable data for evaluating the effects of reforms. CCSSO and WCER recognized the possibility of applying research-based models and instruments for studying curriculum to broader purposes of reporting indicators of curriculum and instruction that could be used by policy-makers and educators. CCSSO submitted a proposal to NSF to lead a study of change in curriculum and instruction related to state standards and state initiatives for improvement of mathematics and science.

State Participation. States interested in examining the effects of reform efforts on classroom instruction and gaining knowledge about the development and use of a survey approach to analyzing curriculum were asked to participate in the study. In 1998, 11 states chose to participate, and state specialists in mathematics, science, assessment or evaluation were invited to join the study management team. The states chose a sample of 20 schools at each of two grade levels (e.g., elementary, middle) for the study. Half the schools selected had high involvement in their state's initiative for improving math or science education ("Initiative" schools), and the other half were schools with less involvement but were similar to the first group based on student demographics ("Comparison" schools).

Data Collection. Teacher surveys of classroom practices were the primary method of data collection. Two teachers per grade level and subject were selected by the principal of each school. The method of selecting teachers was left to principals. They were asked to select teachers of math and science that matched the grades of their state assessment (e.g., grade 4 and 8). Basic information was collected about the schools from principals, and a student survey was conducted in one-fourth of the classes for data validation. Ten of the 11 states chose to focus on elementary and middle school instruction, and one focused on middle and high school instruction.

The Survey for a specific grade level and subject included approximately 150 questions covering:

- Instructional Practices, including classroom activities, assessment, influences on curriculum, and use of technology and equipment;
- Subject Content, including curriculum topics taught by expectations for learning;
- Teacher Characteristics, including teacher education, professional development, and teacher reports on school conditions.

Teachers completed the survey individually, and many used their own time outside of school. Teachers were guaranteed confidentiality, and the main incentive was to contribute to their state's study of reform initiatives in math and science education. At the same time, they were assured data would not be used for school accountability or teacher evaluation purposes. In the spring of 1999, CCSSO obtained completed Surveys from a total of 626 teachers across the 11 states (Iowa, Kentucky, Louisiana, Massachusetts, Minnesota, Missouri, North Carolina, Ohio, Pennsylvania, South Carolina, West Virginia).

Selection of schools and teachers for the study in each of the 11 participating states was based on the degree of school involvement in the state math or science reform initiative. The collected data from the sample schools present sufficient numbers of responses to provide meaningful statistics, such as mean and standard deviation, and the numbers allow analysis of the significance of reported differences related to curriculum and instructional practices in "Initiative" vs. "Comparison" schools. The results from the 1999 Survey reported in the following charts are not nationally representative, nor are they necessarily representative of all mathematics and science teaching in schools in the 11 states.

With results from the 1999 study, we give three examples of the use of enacted curriculum data to analyze systemic, standards-based reform:

(a) Descriptive evidence of instructional practices,
(b) Analysis of professional development influences on instruction,
(c) Analyses of policy influence on instruction, focusing on state initiatives and assessment-instruction alignment.

Descriptive Evidence

The SEC instruments are first and foremost a set of descriptive tools, providing teachers, principals, policy-makers, and others with a "snap-shot" description of practice and related information. A key question for policy-makers is, "To what extent is teaching being changed by standards, curriculum frameworks, and assessments?" (i.e., key policy instruments of standards-based reform). The SEC data provide both direct and indirect measures of policy influence on instruction. For example, direct measures are provided based on teacher ratings of the relative influence of policies on their curriculum, including standards, assessments, and preparation. Additionally, teacher reports of professional development activities provide valuable information about the types of activities, and the impact of professional development on teaching. The SEC data can be analyzed to measure the relationship between amount and types of professional development and impact on instructional practice.

Interpreting Data Charts

The Survey results are reported and analyzed using several formats: Item Profiles, Summary Scales, and Content Maps and Graphs.

Item Profiles present data from individual survey questions, grouped by topic and item format (see Chart 2-1). The data are shown in horizontal bar graphs. The mean is indicated by a solid vertical line, and the shaded bar represents responses that are one standard deviation above the mean and one standard deviation below the mean. Generally the responses at the mean and within the bar represent about two-thirds of all responses to a question. The number of teacher responses per group (e.g., middle, elementary) is reported in parentheses (e.g., 104).

Summary Scale is an average score for a group of 5 to 8 questions in the survey centered on a specific concept underlying curriculum or instruction, e.g., active learning in science (see Charts 2-6, 2-7). Scales are formed by purposeful selection of items and statistical analysis of responses to determine scale reliability (e.g., .81 for communicating math understanding). The selected scale items typically cut across different sections of the survey, and items may have different kinds of responses. The scale measures are "standardized scores," meaning the average score for the scale for the whole group of teachers is set equal to 0, and the standard deviation (a measure of variation in responses) for the whole group is 1. Scale score differences would mean that sub-groups of teachers, e.g., elementary vs. middle school teachers, differ on the concept being measured.

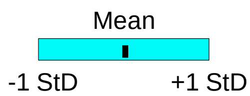
Content Maps and Graphs*.* Teachers report time spent on subject content during the year using a content matrix covering topics and expectations for learning. Responses of teachers are aggregated by grade level and reported with two statistical software programs: a mapping program which gives a three-dimensional picture of variation in time across the whole curriculum (see Chart 2-3), and histograms, which show average percent time by topic and

Influence of Policies on Instruction. Chart 2-1 presents teacher responses to questions regarding factors that influence their science or math instruction. Of the 7 potential policy influences on science, the mean scores for four were in the "little or no influence" to "somewhat positive" range (3 or 4). The mean scores for three of the policy influences were in the range of "somewhat positive" to "strong positive (4 or 5)": (1) District curriculum framework (mean 4.5); (2) State curriculum framework (mean 4.3); and, (3) Preparation of students for next grade or level (mean 4.2). State tests were the next strongest influence reported by teachers with a mean of 4.00.

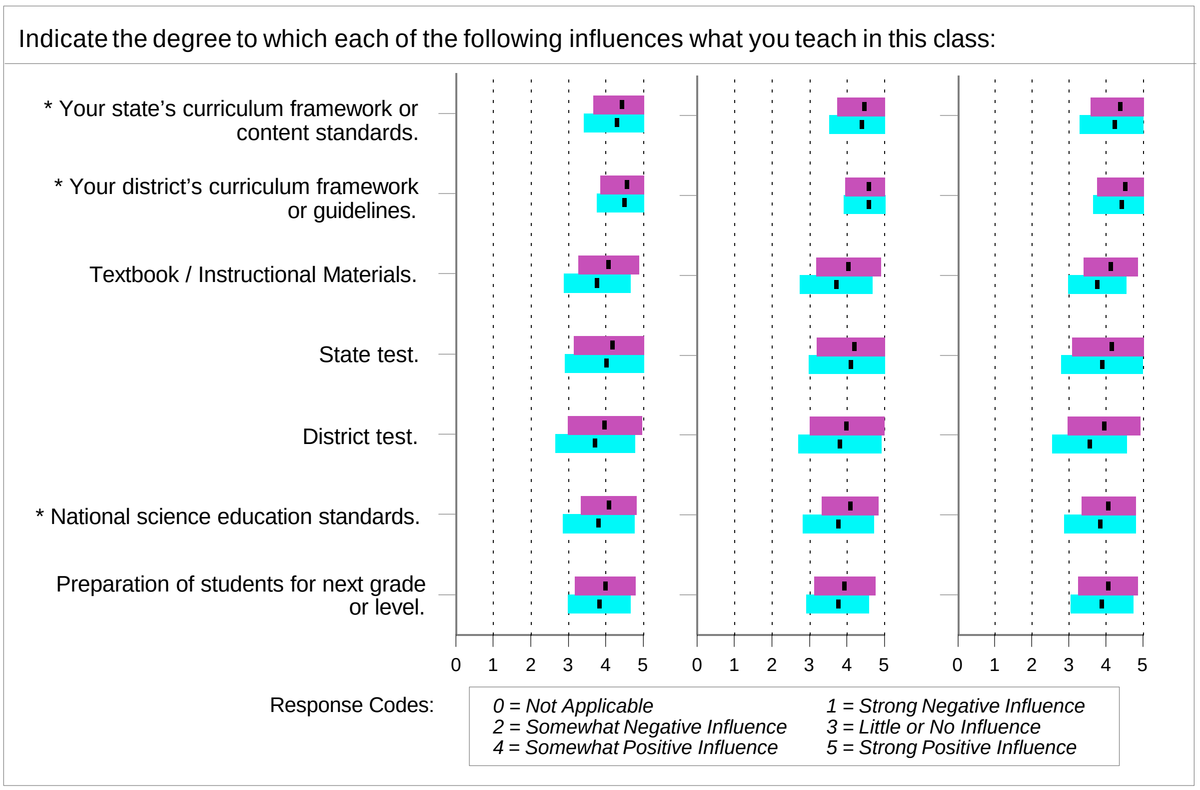
Mathematics teacher responses to the same seven potential policy influences on mathematics instruction are also reported in Chart 2-1. As with science, teachers of mathematics report the strongest influences on instruction were (1) state and district curriculum frameworks and standards, and (2) preparation of students for the next level. Mathematics tests, whether state or district, yield the greatest degree of variability among teachers, with larger numbers of teachers reporting tests as having "little or no influence." Note that all of these results need to be interpreted within the context of the policies of the 11 states and participating districts (see Appendix B).

Chart 2-1
Influences on Instructional Practice in Mathematics & Science

Legend



SEC 11 State Cross State Sample		
Sample (n)	**Elementary** (n)	**Middle School** (n)
Mathematics (300) Science (255)	Mathematics (169) Science (151)	Mathematics (131) Science (104)



* Item included in summary scale.

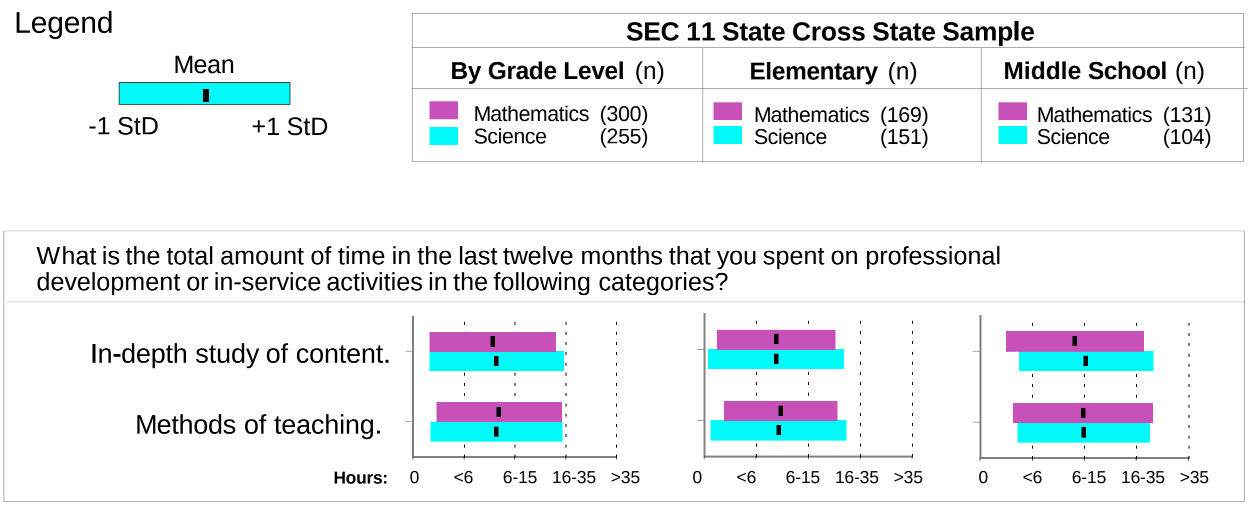
Bordered bar indicates statistically significant mean difference.

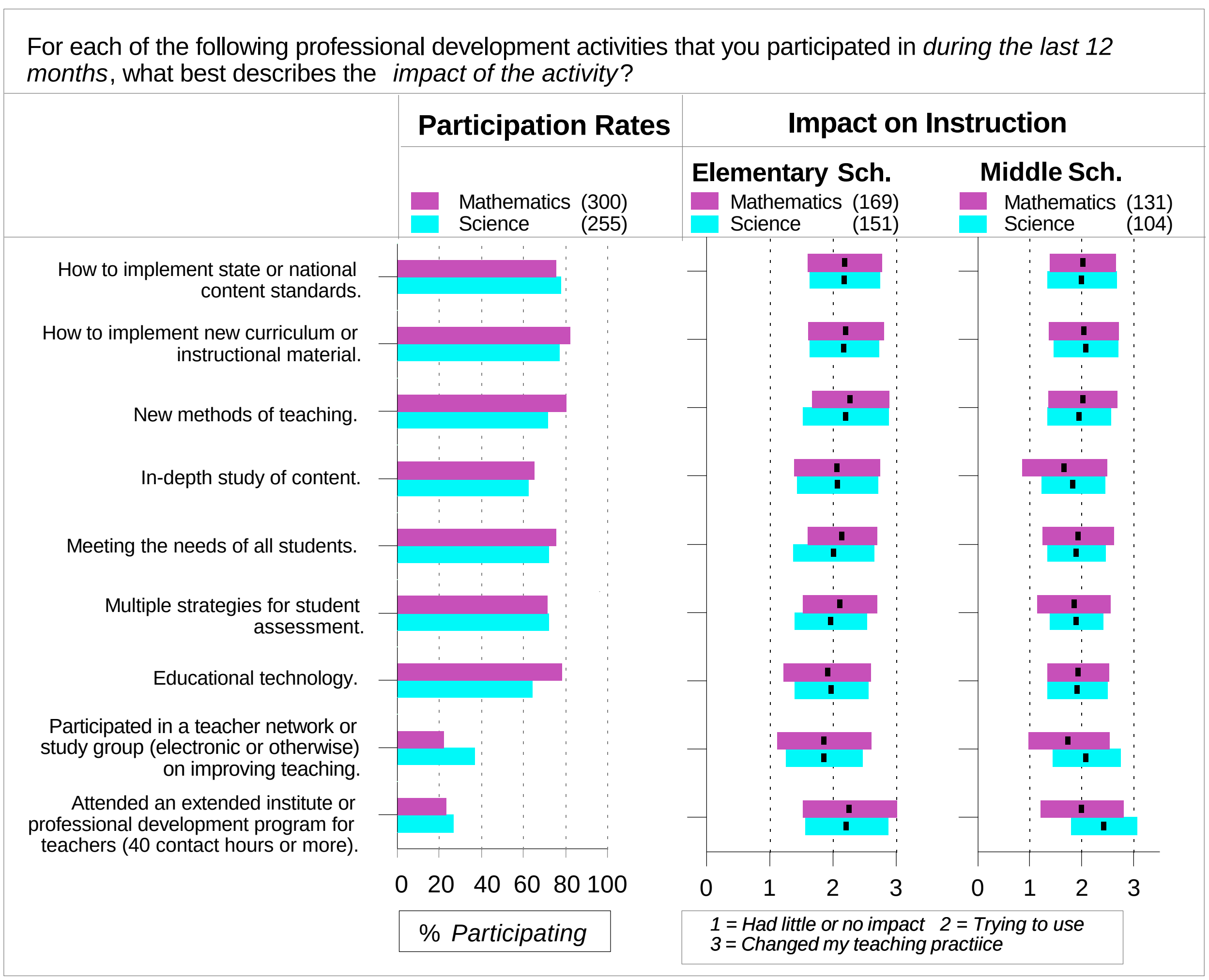
Professional Development Activities. Chart 2-2 illustrates the use of teacher responses on the survey concerning the types of professional development activities attended in science and mathematics respectively. In both subjects, the majority of teachers report attending professional development activities frequently associated with reform initiatives. The 1999 Survey results show that 75 percent of reporting *science* teachers attended professional development activities related to the topic of implementing state or national standards. This was the most commonly reported type of professional development received.

In *mathematics*, slightly more than 80 percent of teachers reported attending professional development activities concerned with implementing a new curriculum or new teaching methods within the past year. Like in science, about three-quarters of the teachers reported attending professional development activities associated with implementing state or national standards, and the use of multiple assessments. Professional development related to educational technology appears more popular among mathematics teachers than science (78% participated in professional development related to educational technology for math, 62% for science).

Both mathematics and science teachers typically reported that they were "trying to use" the information gathered from professional development experiences. About one in five teachers reported that their professional development experiences had caused them to “change their practice” (21% in math, 19% in science, *not shown in chart*).

Chart 2-2
Professional Development in Mathematics & Science





Bordered bar indicates statistically significant mean difference.

Content Mapping. Using the survey data from teachers on curriculum taught, we were able to use a mapping software program to construct 'maps' of the content coverage. Educators have found these graphical representations of content to provide a useful representation of content emphasis in instruction. Content maps can also be constructed for state or local assessments. Content maps can be compared to get a picture of where there is alignment and where there is not alignment between a state assessment and instruction in that state.

Illustrative topographical maps of instructional content are presented for teacher reports from three states on instructional content for Grade 4 mathematics. The sample maps presented in Chart 2-3 indicate that teachers in Massachusetts spend more time on computation (operations by perform procedures) than reported in the other two states. By contrast, Minnesota teachers report more time spent on understanding geometric concepts than reported by teachers from the other two states. Iowa teachers report spending less time on reasoning, solving novel problems, and interpreting topics than the other two states. (See side bar for information on interpreting the content maps.) Chart 2-4 provides a content map for specific concepts taught in the Algebra area of Grade 4 mathematics for the same three studies.

The content maps (Charts 2-3, 2-4) are powerful tools for helping practitioners understand their own instruction and their state assessment. For the purposes of map construction, content emphasis is calculated as though the distinction among topics and the distinctions among cognitive demands are on an ordered, hierarchical scale. However, we note that the topics are not a hierarchy.
We also provide bar graphs of content by expectations (as shown in the example in Chart 2-5, which reports the same information as Chart 2-3), and, for some readers, the graphing report method is more accessible and easier to compare percentages of time across the two dimensions.

Chart 2-3
Instructional Content Map for
Grade 4 Mathematics
(as reported by teachers in three states)

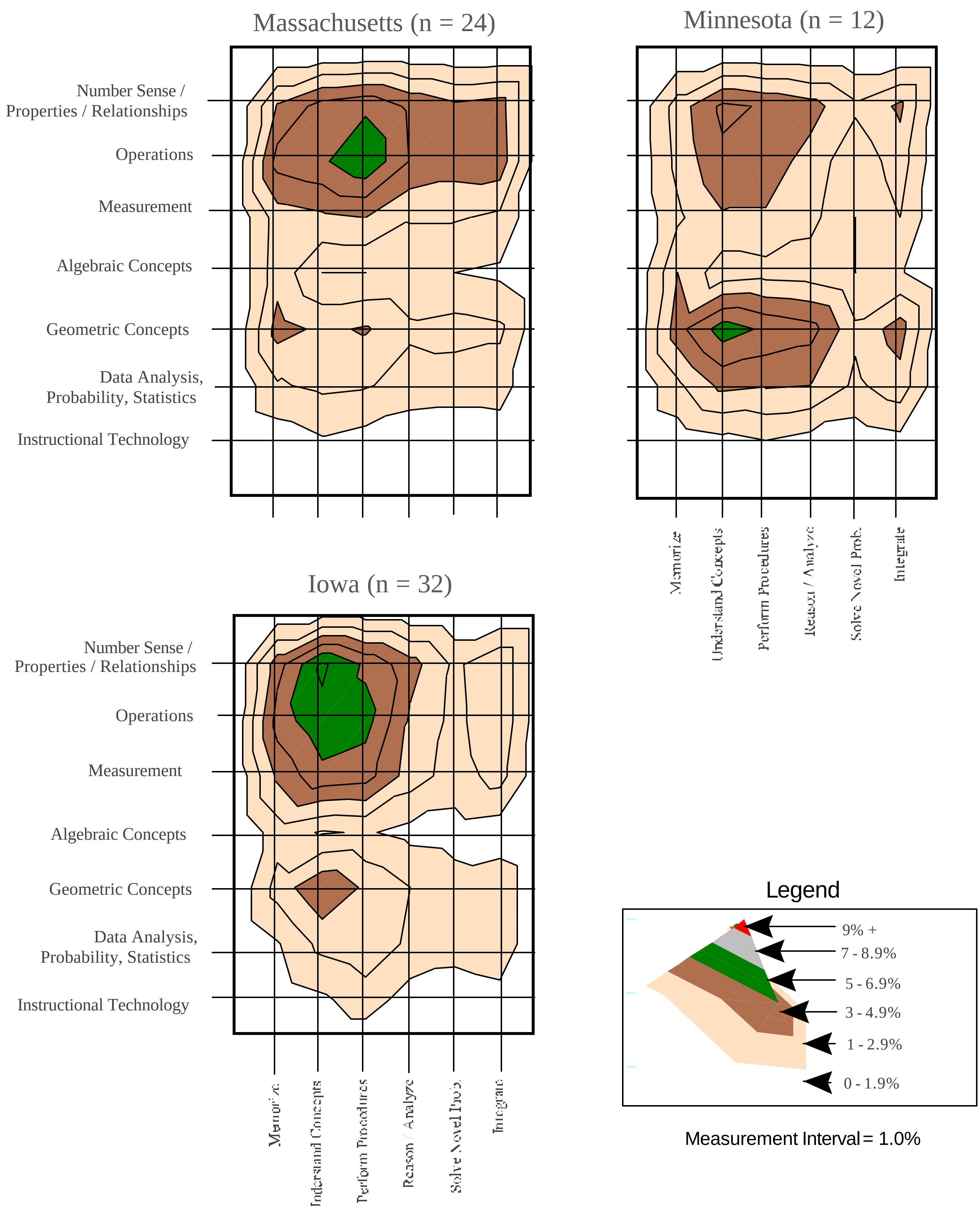


Chart 2-4

Grade 4 Mathematics

Algebraic Concepts

Fine Grain Content Maps

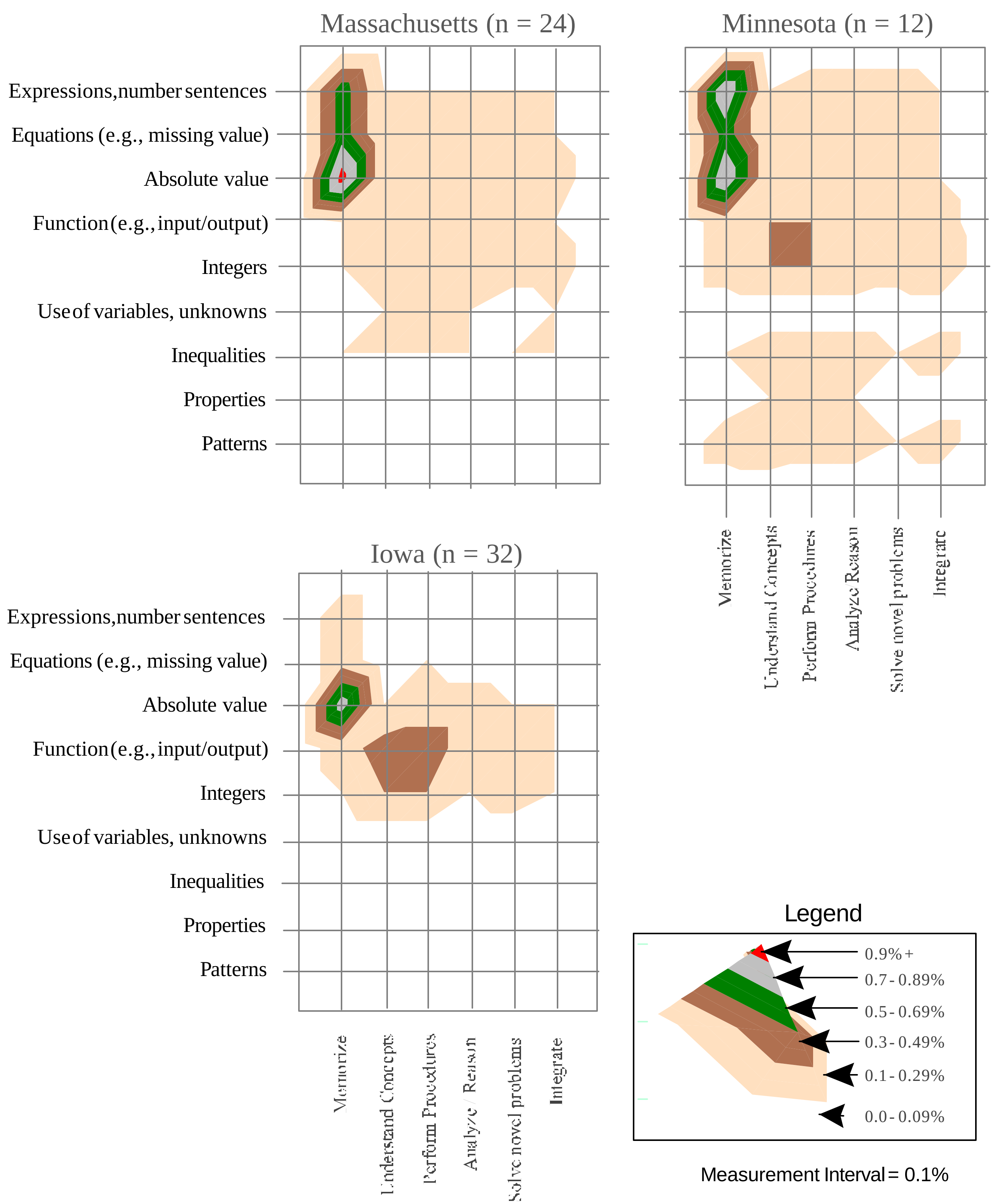
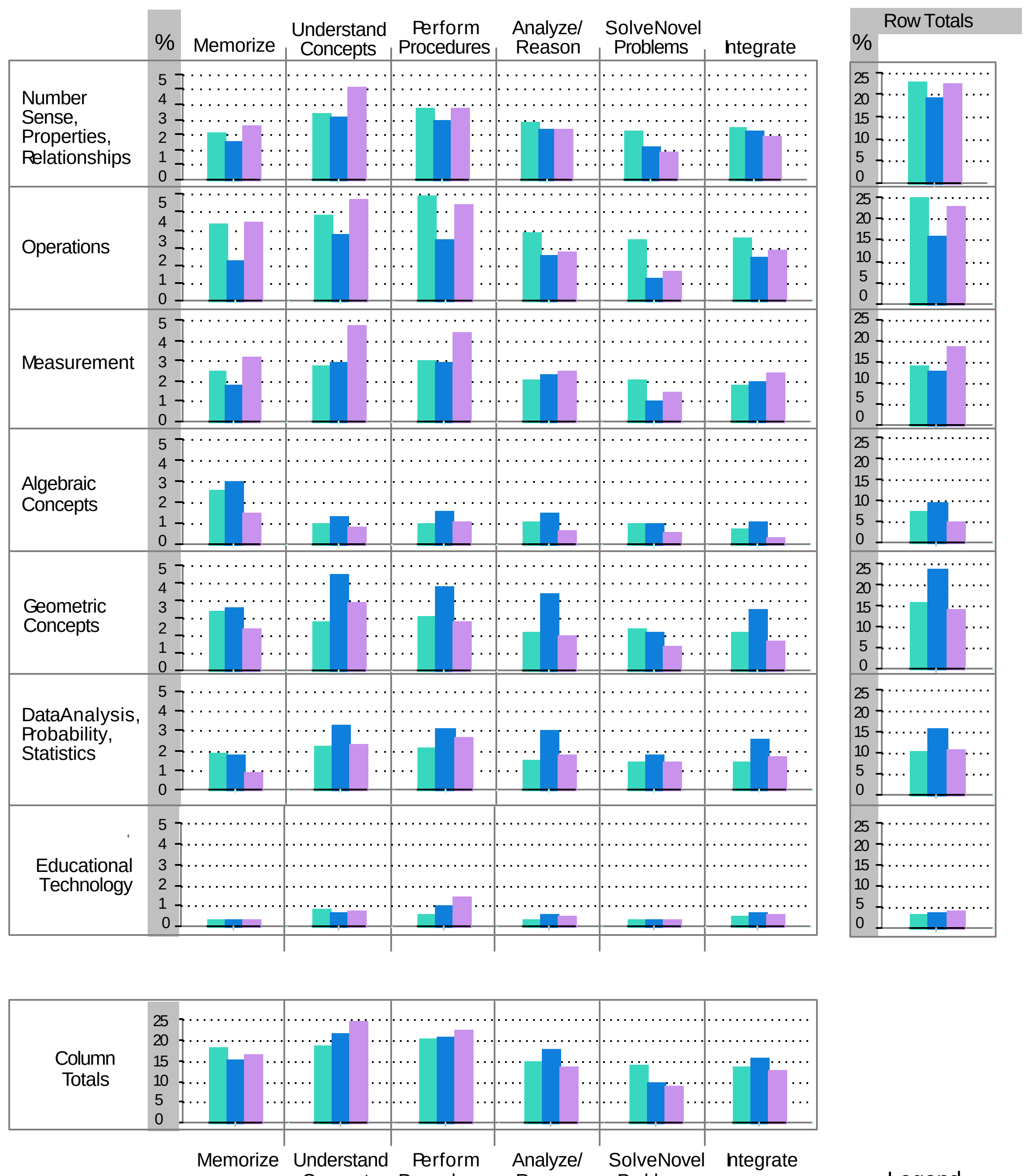


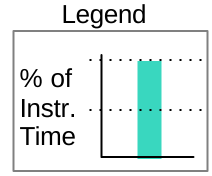
Chart 2-5

Instructional Content Graphs
for Grade 4 Mathematics

(as reported by teachers in three states)



Legend



***Scale Measures of Pedagogy*.** Data in Charts 2-1 and 2-2 suggest that teachers are "getting the message" of the need to use standards as a basis for improvement of teaching in science and math. If this proposition is true, we would expect reports of practice to reflect that influence in some manner. To examine this, we constructed a number of summary scale measures related to instructional practices, and dis-aggregated the data based upon the amount of professional development reported by teachers.

Following data collection, scale measures of central concepts in instructional practices and school conditions for teaching were created, based on math and science standards (see Charts 2-6 and 2-7). Nine scales were created for science, and eight for mathematics. Each scale represents a familiar construct associated with reformed instruction (e.g. communicating mathematical understanding, reasoning and problem-solving, active learning, etc.) or school climate (e.g. professional collegiality, teacher readiness to implement innovative techniques, or provide an equitable environment). An additional scale, "scientific thinking" was added for science. Scales were constructed with 6 to 13 items per scale. (Information on scale construction and scale reliability is reported in Chapter 4, Quality of Data. Items in each scale can be found in Appendix C: Analysis Guide.)

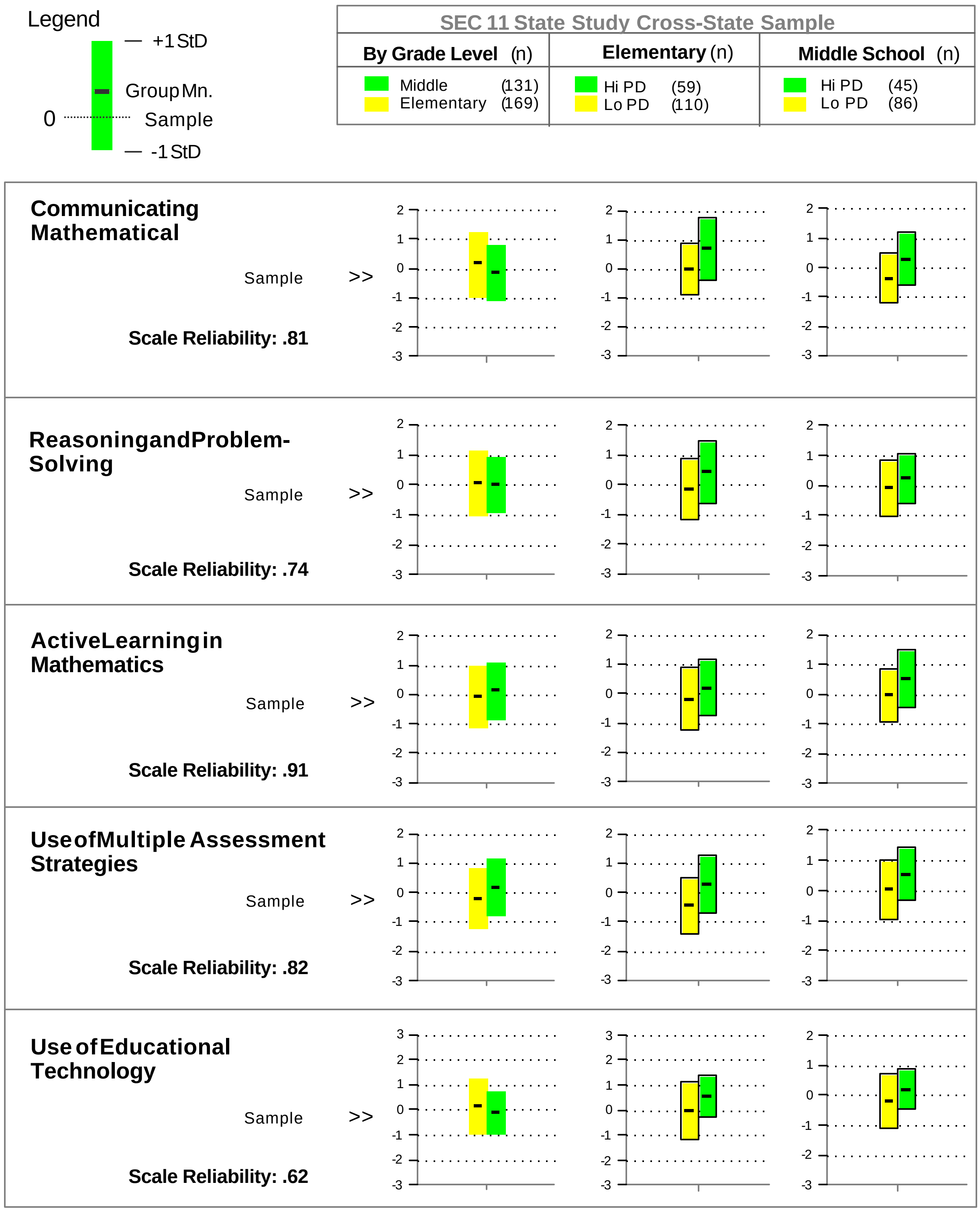
Analysis of Influence of Professional Development on Instruction

To investigate the influence of professional development on teacher practice, respondents were sorted into two groups ("Hi PD" and "Lo PD") based on the amount of professional development in mathematics and science education reported for the previous twelve months. Because elementary teachers generally report less time in professional development activities than middle school teachers, the criteria for assigning teachers to one or another group varied by grade level. Elementary teachers were assigned to the Hi PD group if they reported twelve or more hours of professional development in the areas of math or science education during the previous twelve months. Middle school teachers were assigned to the Hi PD group if they reported twenty or more hours of professional development in math or science education during the previous twelve months.

Chart 2-6 presents the results for mathematics for five scale measures calculated for elementary and middle grades teachers and initiative vs. comparison teachers. Teachers with high levels of professional development report more use of reform oriented practice on each of the five scales reported. Indeed, the mean difference between comparison groups on each of the five pedagogy scales is statistically significant. Chart 2-7 presents the scale results for science. In science, the effects of professional development are most noticeable among elementary teachers. In particular, Student Reflection on Scientific Ideas, Use of Multiple Assessment, and Use of Technology show significant differences between the comparison groups.

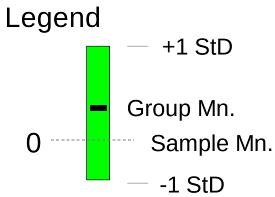
The results presented in these charts indicate that professional development in the subject areas of math and science education are supporting the goals of standards-based initiatives from states. It is possible that the teachers actively involved in and committed to reform are the same teachers that engage in more professional development. To rule out this alternative interpretation, longitudinal data would be needed, as explained in the opening chapter of this report.

Chart 2-6
Mathematics
Scale Measures of Instructional Practices

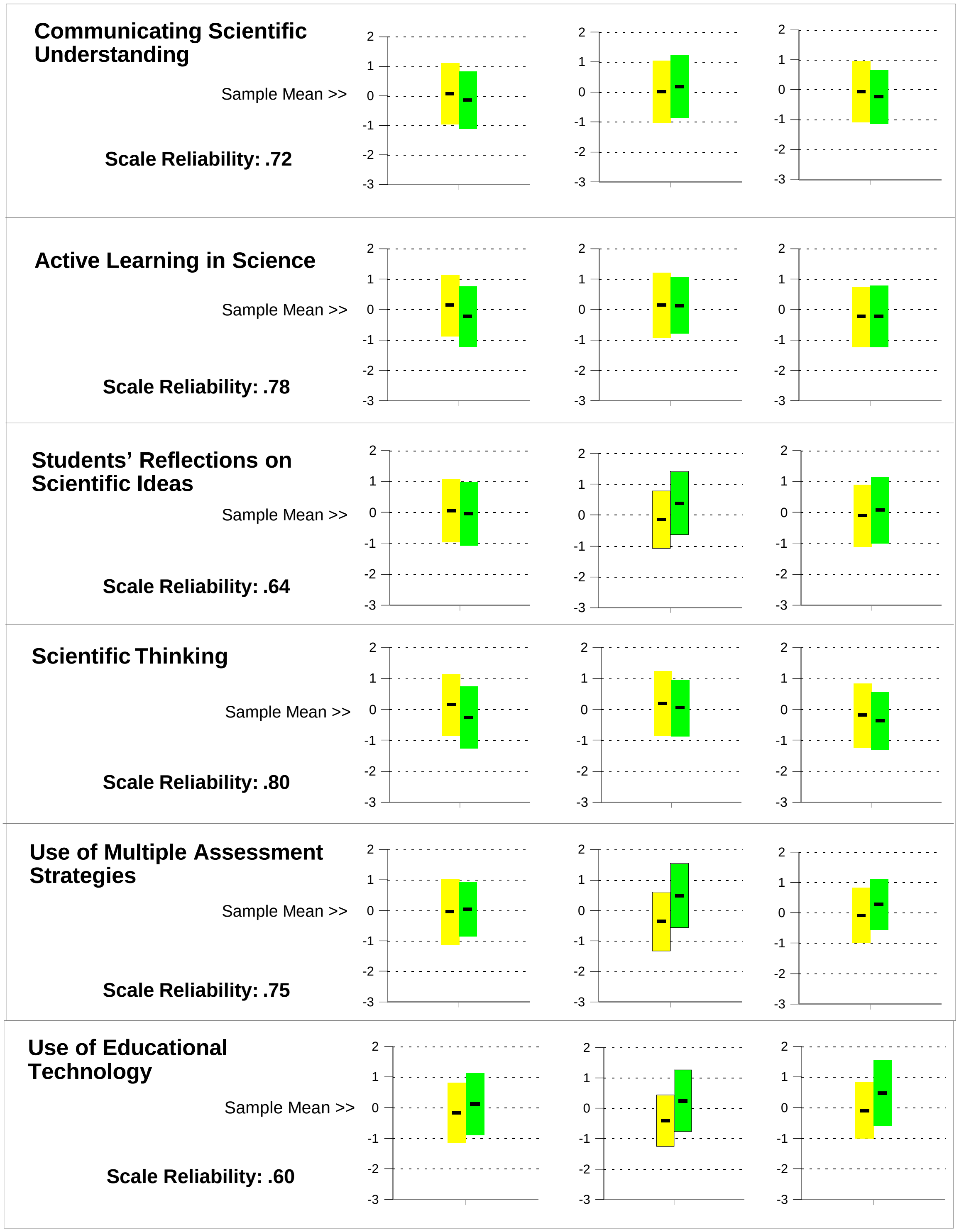


Bordered bar indicates statistically significant mean difference.

Chart 2-7
Science
Scale Measures of Instructional Practices



SEC 11 State Cross State Sample					
By Grade Level (n)		**Elementary** (n)		**Middle School** (n)	
Middle	(104)	Hi PD	(55)	Hi PD	(33)
Elementary	(151)	Lo PD	(96)	Lo PD	(71)



Borderedbar indicatesstatisticallysignificant mean difference.

Analyses of Policy Influences on Instruction: State Initiatives and Assessments

State policy-makers and leaders are interested in having data that informs the question of the degree to which key policies toward standards-based reform (including content standards, curriculum frameworks, and tests) are having the desired effects. The information provided by the Surveys of Enacted Curriculum allows for such analyses. Two general types of analyses are possible through the SEC data:

1. **Comparison of State Initiatives**-- analyses of survey data between states in order to identify distinctive effects of state reform initiatives; and
2. **Alignment Analysis of Assessment and Instruction**--use of procedures developed by Porter and Smithson to calculate a quantitative measure of alignment between instruction and assessment. In previous work (Gamoran, et al., 1997; Porter, 1998), these alignment analysis procedures were used to measure differences in the content of classroom instruction and to predict achievement gains.

Comparison of State Initiatives One method of determining the influence of policies, specifically state policies, is to conduct cross-state comparisons in order to identify the instructional scales or content areas where one or more states show significantly different responses from the total sample. A series of comparisons were made between each state's scale scores and the remainder of the sample. Significance tests were then conducted on the mean differences for the scales scores resulting from these comparisons. Those states with scale measures significantly different from the other states in the sample indicate a difference in the reported practices of teachers that may be attributable to state policies. To make the connection between policy tools and reports of practice would require policy analyses of the sort discussed in Chapter 1. Complete policy analyses to produce quantitative variables for policy differences were not conducted as part of this study, and thus, the results reported below are not sufficient to produce definitives findings on policy impact. Nonetheless, results like these are an important element to such analyses.

The Iowa Case. Of the eleven states participating in the study, only one state (Iowa) had neither a state-administered assessment nor state content standards. Throughout the recent trend toward state standards and assessments, Iowa has maintained its commitment to local control. While the state has recently required districts to establish their own standards and assessment instruments, it is safe to say that Iowa is relatively new to standards-based reform. If other states in the sample have made greater efforts at developing and implementing math and science reforms through policies such as standards, curricula, and statewide assessments and accountability reporting, and these efforts have affected instructional practice, then it would seem reasonable to expect teacher data from Iowa to be less consistent with standards-based concepts reform-oriented on many of the scale measures constructed from the SEC data set. Results of the state-comparisons confirm this expectation for mathematics, however reports from science teachers in the state are similar to the sample of teachers from other states.

Table 2-1 reports results of comparisons between reporting Iowa teachers and the sample of teachers from other states regarding mathematics instruction using the ten scales constructed from the SEC data. For five of the ten scales constructed for mathematics, Iowa teacher reports were significantly different from the rest of the sample. As compared to the broader sample, Iowa teacher data show:

- less teaching of communication for mathematical understanding,
- less use of active learning,
- less teacher preparedness to provide an equitable environment,
- less use of multiple assessments, and
- less influence of standards on instruction.[1]

It should be noted that Iowa consistently ranks high on national standardized mathematics achievement tests such as the NAEP and ITBS. The low measures reported by Iowa teachers should not be interpreted as an indication of low quality instruction in the state, but rather as an indication of less effects of state standards and state systemic initiatives and policies that have been employed by the other states in the sample.

Table 2-1
Mathematics Instruction–Iowa Compared to Other States

SCALE	OUTLIER STATE	State Mean (Std. Dev.)	Other States Mn. (Std. Dev.)
Communicating Mathematical Understanding	Iowa	-.267 (.918)	.009 (1.01)
Readiness for Equity	Iowa	-.410 (1.21)	.125 (.934)
Multiple Use of Assessments	Iowa	-.663 (1.03)	.156 (.942)
Influence of Standards	Iowa	-.273 (.891)	.007 (1.01)

Interestingly, reports from science teachers in Iowa were not significantly different from the rest of the sample when comparisons were made on the eleven scales constructed for science. This may be related to the fact that Iowa has been a participating state in a multi-state science collaborative (SCASS) for nearly ten years, and the schools and teachers active in the project were selected as "initiative" science schools. Also, Iowa has worked within its framework of local control to provide resources and materials for science teachers that reflect and promote the national science standards. Iowa has only recently begun efforts at changing mathematics instruction in the state, again within the framework of local control. Alternatively, across the state samples, science teachers may have focused less on standards-based science reform than mathematics teachers.

[1] See Appendix C for a listing of the items used to construct each scale. See Chapter 4 for information on scale construction and reliability scores.

Other State Reform Differences. In addition to the results for Iowa, a handful of the participating states reported significantly more reform-oriented instruction than the rest of the sample on the scale measures constructed from the SEC data. Table 2-2 displays the results for ***positive outliers in mathematics***. In mathematics, the sample of North Carolina math teachers reported significantly higher results on three scales, compared to the rest of the sample: Influence of Standards, Use of Multiple Assessments, and Teacher Readiness to Provide an Equitable Environment. Other notable differences are listed below.

- The sample results from Minnesota teachers stand out from the rest of the sample on Teacher Readiness for Innovative Practice and Professional Collegiality.
- Sample data from Massachusetts and Louisiana teachers indicate that teachers spend significantly more time on Communicating Mathematical Understanding.
- North Carolina teachers reported significantly more influence of standards on instruction than the other states.

Table 2-2
Mathematics Instruction–State Differences

SCALE	OUTLIER STATE	State Mean (Std. Dev.)	Other States Mn (Std. Dev.)
Influence of Standards	North Carolina	**.520** (1.12)	**.007** (.961)
Use of Multiple Assessments	North Carolina	**.574** (.799)	**-.008** (1.00)
Readiness for Equity	North Carolina	**.564** (.956)	**-.008** (.981)
Readiness for Innovation	Minnesota	**.632** (.792)	**-.004** (.999)
Professional Collegiality	Minnesota	**.845** (1.12)	**-.005** (.967)
Communicating Mathematical Understanding	Massachusetts Louisiana	**.597** (1.14) **.437** (1.15)	**-.108** (.935) **-.119** (.923)

Table 2-3 highlights several examples of positive outliers in science.

- Massachusetts teacher data stood out from the sample with significant differences as measured by three scales: Teacher Readiness to Provide an Equitable Environment, Student Reflection on Science, and Multiple Use of Assessments.
- Both Louisiana and West Virginia teachers reported significantly more Use of Lab Equipment and Educational Technology.
- Minnesota teachers reported more teaching for Communicating Scientific Understanding, and Kentucky teachers reported more Professional Collegiality with other science teachers.

Table 2-3
Science Instruction–State Differences

SCALE	OUTLIER STATE	State Mean (Std. Dev.)	Other States Mn. (Std. Dev.)
Teacher Readiness for Equity	Massachusetts	**.477** (.945)	**-.009** (.986)
Student Reflection on Science	Massachusetts	**.306** (.857)	**-.006** (1.02)
Multiple Use of Assessments	Massachusetts	**.388** (.808)	**-.007** (1.02)
Use of Educational Technology	Louisiana West Virginia	**.446** (.785) **.769** (1.11)	**-.005** (1.01) **-.007** (.959)
Communicating Scientific Understanding	Minnesota	**.502** (1.39)	**-.004** (.949)
Professional Collegiality	Kentucky	**.501** (.780)	**-.006** (1.01)

These comparisons are suggestive of the varied impact of reform across various states in the study, but alone are insufficient to make a strong argument. Such results would need to be combined with analyses of policy that demonstrated characteristics of consistency, prescriptiveness, power, and authority (see Chapter 1) in order to explain why one state might look different from another on these scale measures.

Alignment Analysis of Assessments and Instruction. A new and potentially powerful technique for analyzing the relationship between instruction and policy instruments (most notably state assessments) is made possible by the content languages developed for describing and quantifying instructional content used in the Surveys of Enacted Curriculum. (The findings in this section of the report are also described in a recent article, by Porter and Smithson, "Are Content Standards Being Implemented in the Classroom?"(2001).)

By utilizing the same language to analyze the content of assessment instruments, it is possible to compare the two data sources (instruction and assessment) and calculate an index measure that functions similarly to a correlation. A measure of '1' represents perfect alignment between instruction and assessment, while a measure of '0' indicates no intersection of content of instruction and content of student assessment instrument.

Methodology. To conduct an alignment analysis, information is needed on the relevant 'target' for instruction (i.e. state, national, or other achievement test). Chart 2-8 presents illustrative results for Grade 8 Science in State ‘B’. In this example, teacher reports from State ‘B’ are compared to content analyses of the Grade 8 science test used for assessment and accountability in the state, as well as to content analyses of the Grade 8 Science NAEP assessment. The example reveals that teachers in State ‘B’ spread their instruction over more content areas than are tested. Note that the state assessment had few, if any, items associated with measurement and calculation in science, while teachers report content coverage in this area. Alignment analyses reveal that instructional content was not very well aligned with either the state test or the NAEP test for Grade 8 science (.17 for the state test, and .18 for the NAEP test).

Of the 11 states included in the study, six states participated in a sub-study to analyze the content of their assessments. The assessments analyzed were mathematics and science tests in grades 3, 4, or 5, at the elementary school level, and grades 7 or 8 at the middle school level. For some states, multiple forms were analyzed. All grades 4 and 8 NAEP items were also content analyzed.

Tests were content analyzed, item by item, using the same language and distinctions for describing content (topics by cognitive demand) as employed in the survey. Content analyses were conducted during a two-day workshop in the summer of 1999. The analysis teams were comprised of six state mathematics specialists, six state science specialists, three math educators from universities, and four science educators from universities and research organizations.

Teachers described the content of their instruction using the SEC content instruments item examples. They reported the amount of time spent in instruction over the past year on each of several topics. For each topic taught, they reported the degree to which one of several expectations for students were emphasized, including memorize, perform procedures, solve novel problems, and apply information.

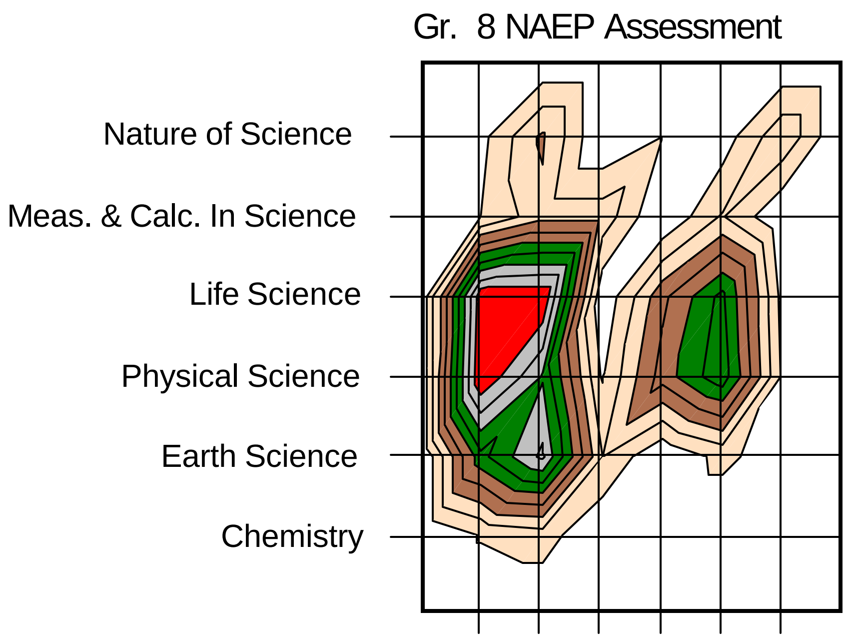
Findings on Alignment. Three types of analyses can be demonstrated with the data:

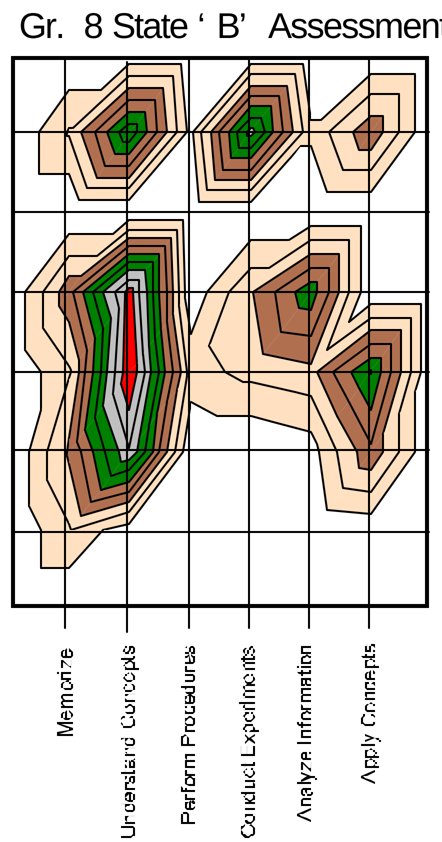
1. Assessment-to-assessment alignment, including state assessments with NAEP,
2. Instruction-to-assessment alignment,
3. Instruction-to-instruction alignment.[2]

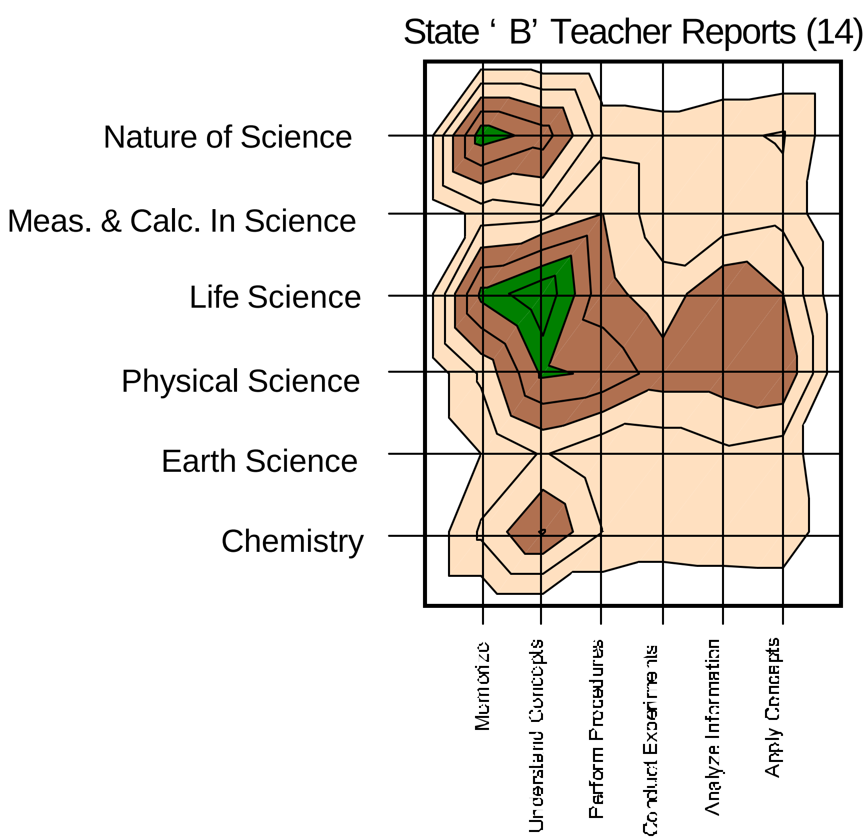
---

[2] The third type of alignment calculated for the study looks at alignment of instruction across states. The measure speaks primarily to the sample of teachers represented in the study, and is discussed in Chapter 4.

Chart 2-8
Grade 8 Science
Alignment Analysis

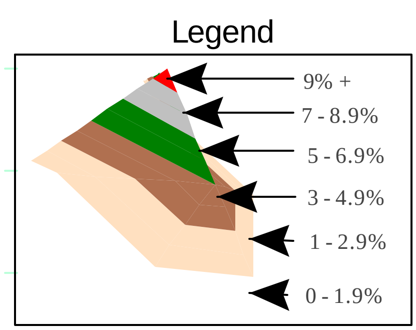






Alignment between Assessment & Teacher Reports of Practice:

Instr.	To State Test	**.17**
Instr.	To NAEP Test	**.18**



Measurement Interval = 1.0%

Calculation of ***assessment-to-assessment alignment*** permits examination of how similar or dissimilar state tests are to one another and to NAEP assessments. The results can be displayed in a manner similar to a correlation matrix (see Table 2-4) and are similar in interpretation. The main diagonal of the matrix indicates comparison of a state test to itself, and thus has an alignment measure of 1.00. The off-diagonals report alignment between one state and another. The closer to 1.00 a state's test is to another, the more similar those two assessments are. The off-diagonal measures in Table 2-4 suggest that despite some overlap, more that half of the items on a given state's assessment test content not assessed by other states.

Table 2-4
State to State Alignment of Grade 4 Mathematics Assessments

	STATE A	STATE B	STATE C	STATE D	STATE E	NAEP
STATE A	1.00					
STATE B	0.41	1.00				
STATE C	0.37	0.34	1.00			
STATE D	0.41	0.41	0.45	1.00		
STATE E	0.44	0.43	0.36	0.44	1.00	
NAEP	0.39	0.39	0.26	0.36	0.37	1.00

Summary measures for the two grade levels[3] and two subjects surveyed are presented in Table 2-5. The summary measures indicate that in at least five states with state assessments, instruction tends to be slightly more aligned with the state test than with the NAEP, regardless of subject or grade level.

Table 2-5
Average Assessment Alignments Across States

	Average State to State Alignment	Average NAEP to State Alignment
Grade 4 Math	0.41	0.35
Grade 8 Math	0.33	0.30
Grade 4 Science	0.33	0.29
Grade 8 Science	0.28	0.20

The degree of alignment ***between state assessments and instruction reported by teachers*** was also calculated for each state and compared across states (see Table 2-6). To the extent that a test is a target for standards-based reform, and to the extent that standards-based reform is having an effect, alignment

[3] For most states, grades 4 and 8 tests were analyzed. However, in some states, grade 3 and grade 7 tests were used, as these were the grades in which mathematics or science was tested in that state.

of instruction in a state should be higher for that state's test than for the test of another state. The desired level or degree of alignment between instruction and assessment (i.e., how close to 1.0) is not easy to determine, and it is not a statistical issue. The degree of alignment is a policy issue for a state and its educators. As a practical matter, perfect alignment is not achievable, nor, as a policy matter, is it desirable. An assessment can at best only sample the scope of knowledge and skills we we wish students to learn. Precisely what measure of alignment is most desirable is therefore a difficult question to answer. Nonetheless, alignment analyses such as these serve to provide policy makers a basis for discussions about the extent assessment are and should be "aligned" to instruction.

Table 2-6
Average Among States: Instruction to Assessment Alignment

Subject / Grade	State Instruction to State Assessment	State Instr. to Other Assessments	Avg. Instruction to NAEP Assess.
Grade 4 Math	0.42	0.33	0.41
Grade 8 Math	0.33	0.24	0.22
Grade 4 Science	0.37	0.28	0.23
Grade 8 Science	0.23	0.23	0.14

In Table 2-6, we can observe that instruction in a state did tend to be more aligned to that state's test than to the tests of other states, suggesting that standards-based reform is bringing instruction into alignment with state tests. With the exception of Grade 4 mathematics, alignment of instruction to assessments was higher for state tests than for NAEP tests. Instruction tended to be least aligned to Grade 8 NAEP science (.14). To the extent that one views NAEP assessments as being oriented towards mathematics and science standards, Grade 4 mathematics instruction appears to show the most evidence for standards-based instruction.

Three caveats are necessary. First, regarding the extent that a state test is or is not aligned to a state's content standards, one might not want instruction to be tightly aligned to the state test. Nonetheless, to the extent that a state test is used in an accountability program, it may have an influence over instructional practice. Second, these data are illustrative only. The samples of instruction in each state cannot be taken as representative of that state, as the sample was neither randomly selected nor sufficient in size for generalization. (An example of analysis of one state's instruction to assessment alignment is given in Chapter 3.) Third, the data are not longitudinal. The purpose here is to illustrate the types of analyses possible.

While achievement data were not collected as part of this study, the SEC tools also allow for investigation of the relationship between achievement scores and instruction. Assuming one has item level data on student performance, as well as content analyses of each item, achievement results can be arrayed into the content language and portrayed as a "map" and/or used to calculate a measure of alignment between instruction and achievement gains.

Trend Data

The data from the 1999 Survey suggest the influences of reforms on practice, but this kind of "snap-shot" report from a single point in time provides only limited evidence of effects. We would prefer to analyze change over time in the direction of reform. For the participating states, these results present a modest baseline (depending on the representativeness of the sample for a given state) for monitoring the progress of reform in those states. Ideally, we recommend that enacted curriculum data be collected every few years from a representative sample in order to track changes in reports of teacher practice as initiatives have more time to affect practice.

Longitudinal data were not collected as a part of the 11-state study, but this approach is being used in a new experimental-design study of the effects of using enacted curricula data as a tool for improving instruction. With support from NSF, the three-year study titled “Improving Effectiveness of Instruction in Mathematics and Science with Data on Enacted Curriculum” (REC-0087562) began in 2000. It is a collaborative project of the Council of Chief State School Officers, the Wisconsin Center for Education Research, and the Regional Alliance for Mathematics and Science Education Reform at TERC.

Chapter 3: Use of SEC Data in Schools

Teachers, schools, and districts seek ways to improve dialogue among teachers regarding strategies for improving instruction and content taught in classrooms. The data and analyses from Surveys of Enacted Curriculum can help with both the process and the substance of interaction among teachers regarding instruction in math and science education. These data can be useful in both informal, small-group discussions among teachers as well as formal professional development programs.

Framing the Issue

Teachers often reported they have little time or opportunity for working with colleagues in their schools to improve teaching. Our survey indicated that teachers said they would benefit from more opportunities to work with other teachers (CCSSO, SEC Survey results, 2000). Schools found it difficult to create additional time in the school year or the school day for teacher planning and work with colleagues, even though this was the desire of both teachers and administrators. Additionally, many teachers were not taught to work collaboratively with their colleagues. The organization and schedule of schools did not promote teaming with other teachers, and many teachers did not feel comfortable sharing strategies and methods of teaching with colleagues.

Current Professional Development. We also know from teacher surveys that much of the time that was spent in formal in-service education or professional development did not focus on the curriculum or subject content they are expected to teach (Porter, et al., 1999). In-service activities in schools or districts may cut across subjects and not directly address the curriculum taught by teachers, e.g., when the focus is on topics like use of technology or discipline. On the other hand, professional development for recertification or renewal of teacher licenses was typically based on course credits or CEUs. Formal courses for license renewal took teachers away from their school to a university or PD center where they worked with teachers from other schools or districts.

Data to Focus Improvement. The SEC can help to address these problems by assisting schools and staff to focus their formal professional development experiences and teacher networking on curriculum content and practices used in school. The SEC is designed to be given to all teachers in a school, not to representative samples. The data analysis phase should involve teachers and administrators in reviewing data and making interpretations about the meaning of findings. Results should be reported at the school level where educators can use the data. But the results will not be useful if surveys are treated as accountability tools, because they rely on teacher self-reports about their instructional practices. If teachers are cognizant of being judged on what they report, they are likely to bias the data about their activities.

The SEC is designed to cover the range of classroom instructional practices and content that might be taught or used. The questions are designed to be neutral. There are not right and wrong answers. The goal is for teachers to reflectively review their own practices in relation to those of other teachers. Teachers

should be able to look at data about practices across their school or district without worry about criticism from others.
Curriculum data at the school level can provide useful comparisons by which teachers can:

(a) examine practices used by different teachers,
(b) relate the range of practices to the system or school goals for learning, and
(c) consider his/her own practices in relation to others and the goals for learning.

The data on classroom practices can also be statistically analyzed with achievement scores to see if there are patterns in student outcomes in relation to practices.

Analyses of the data from the 1999 Study reported in Chapter 2, as well as earlier reports, used the cross-state sample of over 600 teachers as well as state samples (CCSSO/WCER, 2000). In this chapter, we demonstrate two ways to analyze SEC data with greater detail for use by schools and teachers by providing examples of:

(a) how the Survey data can be analyzed and displayed at the school and teacher levels;
(b) how the data can be analyzed to examine content taught within a topic area. The goal is to model analytic strategies that local specialists and professional development leaders can use in applying data for math and science improvement.

Using Enacted Curriculum Data within a School

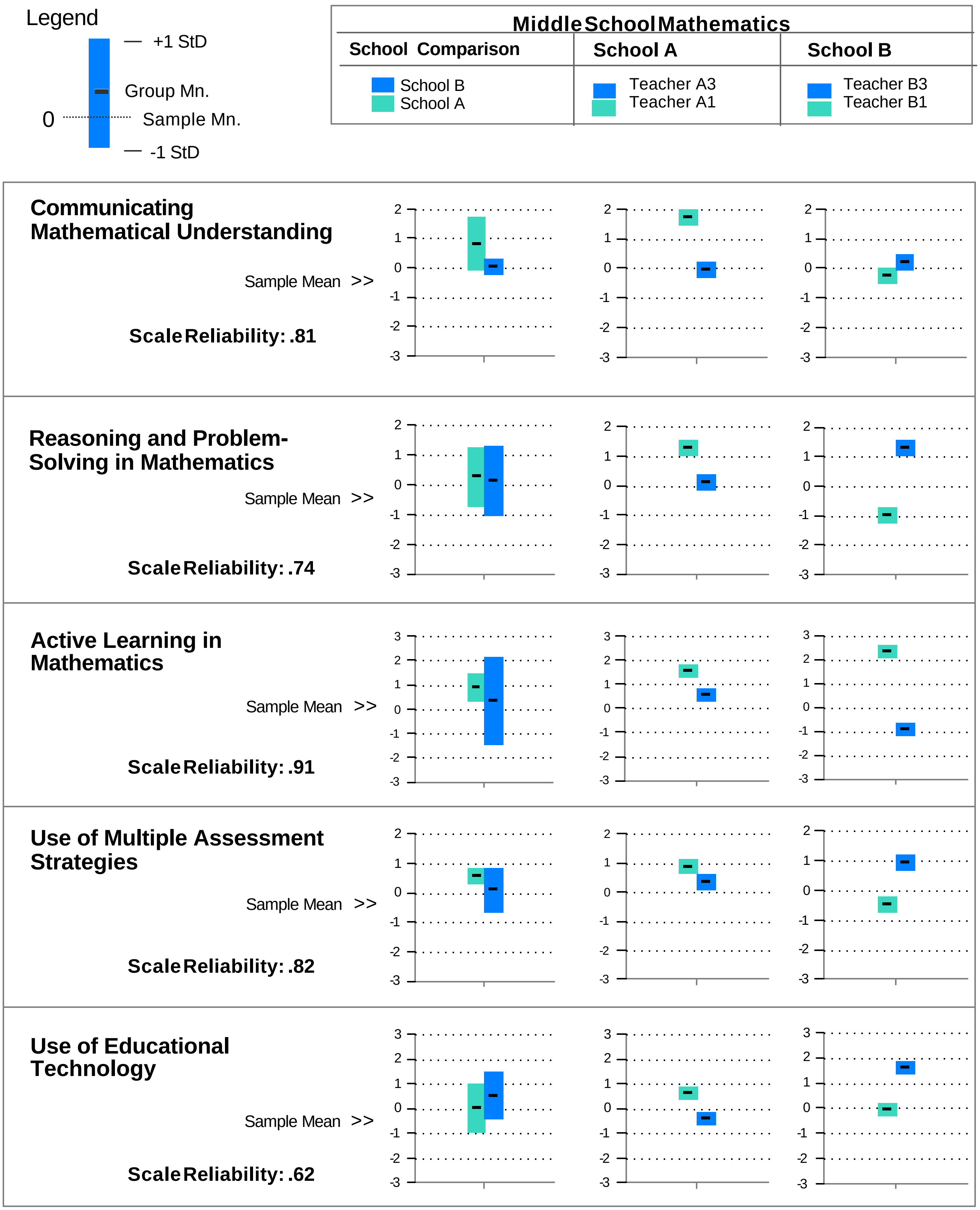
The following pages include selected data charts from responses of Middle Grades Mathematics teachers and Middle Grades Science teachers in the 1999 Study. We provide examples of how data can be reported to show averages and range of responses within schools among teachers. In the following example descriptions of SEC data reported at the schools, we also identify some of the skills that educators will gain through their own analyses of data. We outline four typical steps in working with teachers to analyze and apply enacted curriculum data:

1. Reading and interpreting charts,
2. Examining differences in practices,
3. Comparing instruction to standards,
4. Conduct in-depth discussions and collaboration.

Gain Familiarity in Reading and Interpreting Data Charts. The data on enacted curriculum are generally reported with three different formats (as described in Chapter 2). Teachers and others involved in analysis will need to understand the statistics and how they are represented in the data reporting charts.

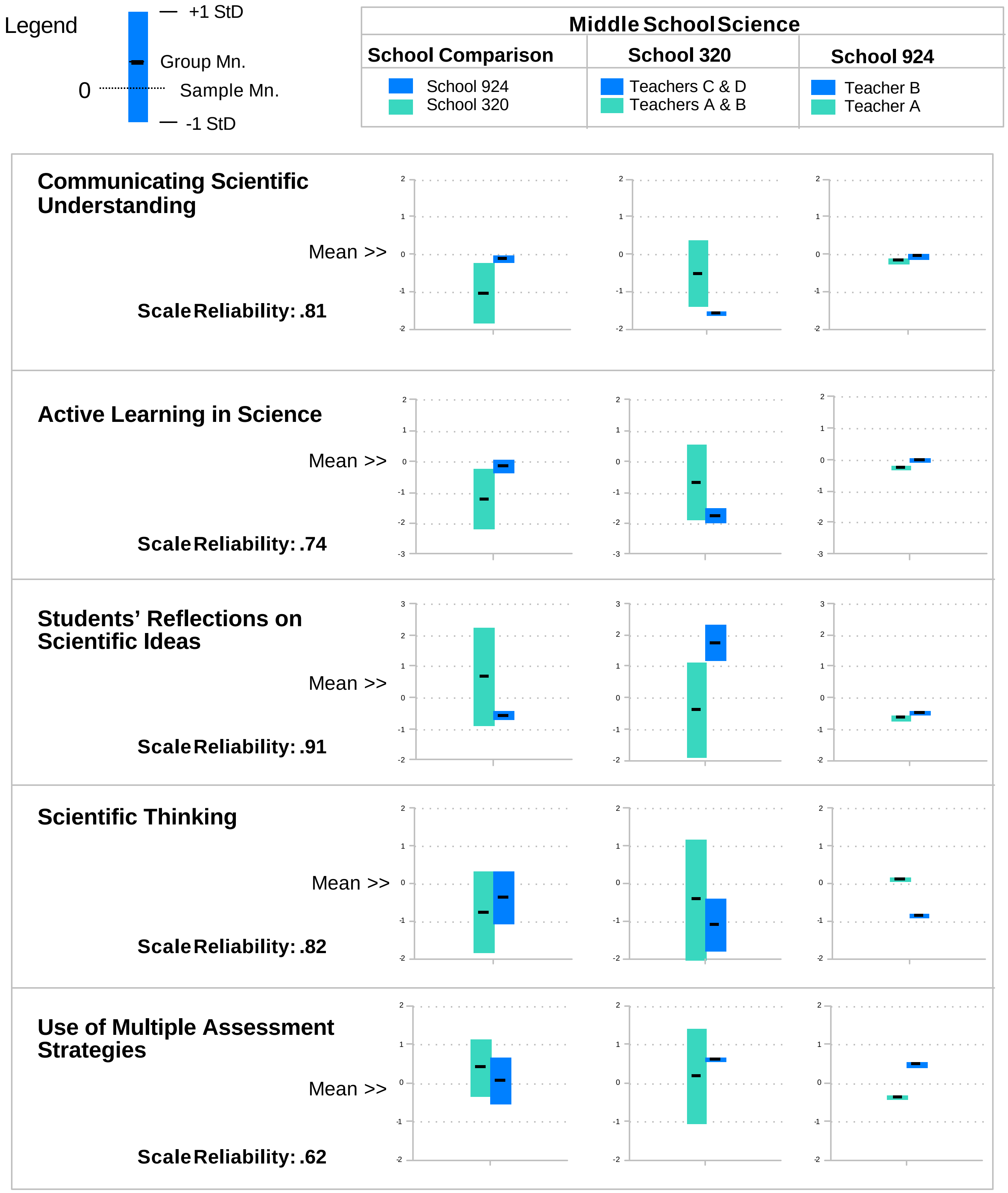
The *Scale Measures* in Charts 3-1 and 3-2 provide average teacher responses across several items that together represent an important concept, strategy, or policy related to instruction. In this Survey, mathematics and science standards of states are represented in the scales (Communicating, Reasoning/Problem Solving, Active Learning, Scientific Thinking, etc.). The scales provide comparisons to the average, not a percentage. The average score among all schools and teachers is set at 0, so that specific school scores and teacher responses can be viewed in relation to the overall average. Thus, under Communicating Mathematical Understanding (Chart 3-1), School A is significantly above the average in time spent on students learning how to communicate mathematical understanding, while School B was right at the average. The teachers in School A differed markedly on this scale measure, i.e., Teacher A1 spent a lot time on Communicating, while Teacher A3 was right at the overall average.

Chart 3-1
Middle School Mathematics
Scale Measures of Instructional Practice
State Sub-Sample



Source: CCSSO/WCER, Survey of Enacted Curriculum, 1999

Chart 3-2
Middle School Science
Scale Measures of Instructional Practice
State Sub-Sample



Source: CCSSO/WCER, Survey of Enacted Curriculum, 1999

The *Item Profiles* (Charts 3-3 through 3-6) depict the percent of time teachers reported that their class spent on each type of instructional activity or teaching method. The survey questions were designed to cover a broad range of possible instructional activities in classrooms.

The *Content Maps* (Charts 3-7, 3-9) and *Content Graphs* (Charts 3-8, 3-10) allow comparison of the subject content taught in the same subject among different schools and classrooms. Both methods of displaying survey results on subject content are based on the amount of time teachers report their class spent during one year on content topics by the teacher's expectations for student learning in math or science. The vertical dimension shows main content topics, and the horizontal dimension shows expectations for learning.

Examine Main Differences in Practices Among Schools and Teachers. After teachers are comfortable with reading the data charts, one approach to analyzing school and classroom differences is to focus on some of the main differences in instructional activities.

Chart 3-3 *(Instructional Activities in Mathematics)* summarizes and compares responses from teachers in two middle schools on a variety of instructional practices. This chart allows us to look across the range of activities and identify similarities and differences among schools and teachers. For example, in the first column, we see that Schools A and B differ on teacher demonstrations (or lecture). School A varied from 12 percent to 42 percent of class time (mean 28%), while School B varied from 10 to 28 percent (mean 20%). Hands-on materials were used differently, with School A teachers averaging 10 percent of time, while School B averaged 22 percent of time. Use of computers and calculators differed between Schools A and B, averaging about 18 percent of class time in School B, and 28 percent for School A.

Data on teaching practices, such as those displayed in these charts can be used by teachers to examine and reflect on their own practices, and to ask each other more detailed questions about their methods and the results they see. In Chart 3-3,Teacher A3 used demonstrations (lecture) almost half the time (42%), while her colleague (A1) used them only 10 percent of time. Students with Teacher A3 spent more time taking notes in class. Teacher A1 used portfolios with students, did more activities outside the classroom, and made more use of educational technology. However, both teachers spent similar time on several of the other teaching activities, including working in pairs/small groups, problem solving, collecting/analyzing data, using hands-on materials, and taking tests. Then, for further analysis, educators can examine more detailed data about these activities, such as problem-solving, shown in Chart 3-5.

The data in Chart 3-4 *(Instructional Activities for Science)* shows differences in instruction between two middle schools (320, 924), such as time listening to teacher (e.g., lecture), hands-on lab activities, working in small groups, and performing activities outside class. However, closer inspection of the teacher data reveals that the differences among teachers within their respective schools were greater than the differences between schools. Teachers within both schools differed in their use of lab activities, writing in science, use of computers to learn science, individual work, demonstrations, and time on tests.

In Chart 3-6, we can look more closely at specific types of activities carried out by students during investigations and experiments. The data indicate that all the teachers reported using almost all of these practices, but there were substantial differences in time spent on using equipment, collecting data, designing experiments, and making predictions/hypotheses. These time differences provide indicators of differences in how students are taught to do science.

Chart 3-3
Middle School
Instructional Activities in Mathematics
State Sub-Sample

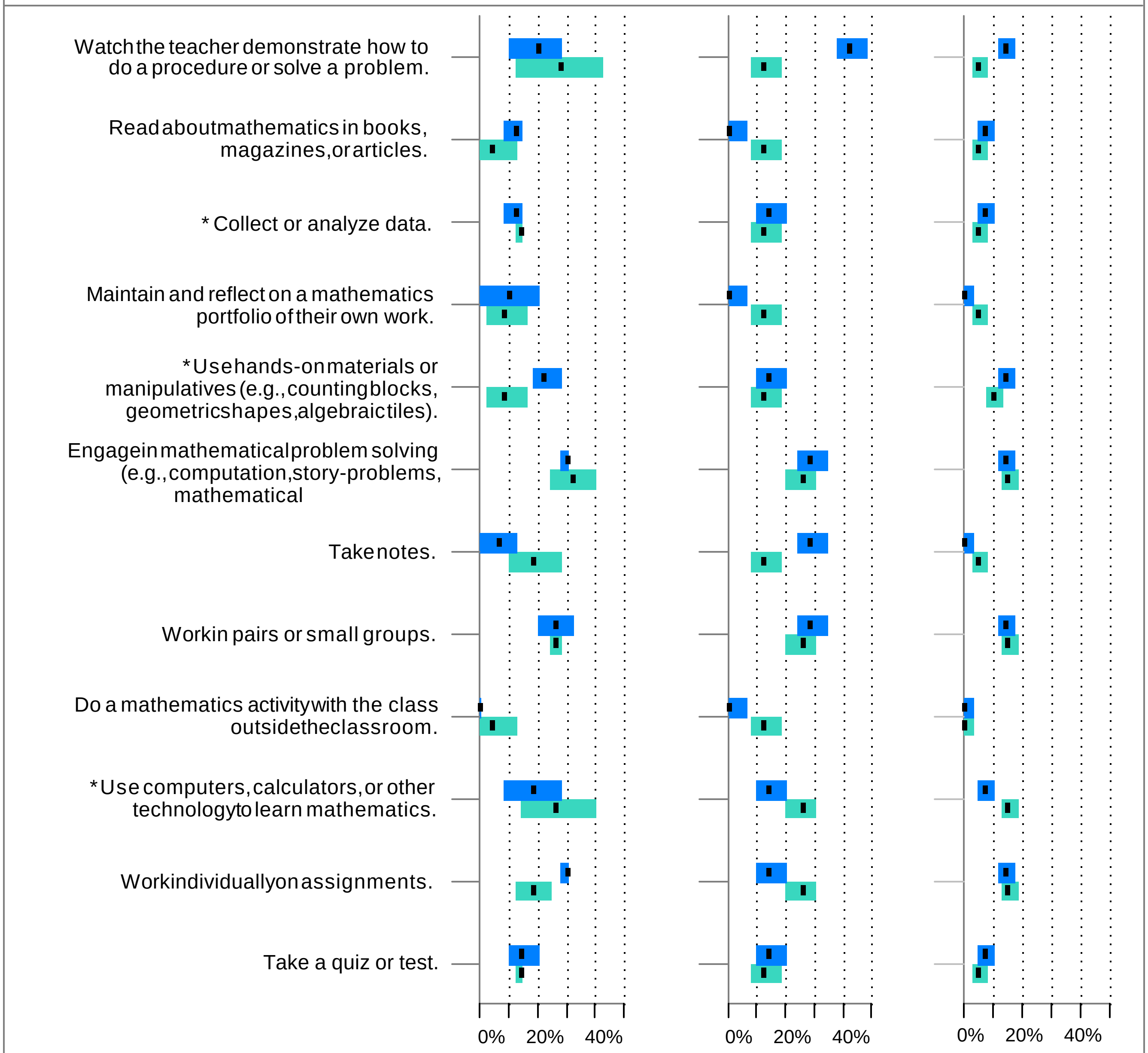
Legend

Mean

-1 StD +1 StD

Middle School Mathematics		
School Comparison	**School A**	**School B**
School B School A	Teacher A3 Teacher A1	Teacher B3 Teacher B1

What percentage of mathematics instructional time do students in this class spend on the following activities?



* Item included in summary scale.

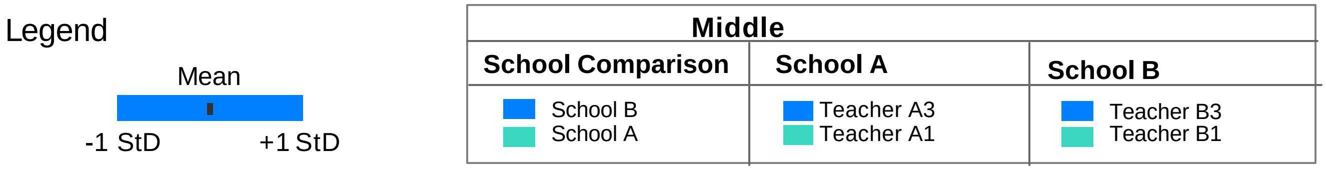
Source: CCSSO/WCER, Survey of Enacted Curriculum, 1999

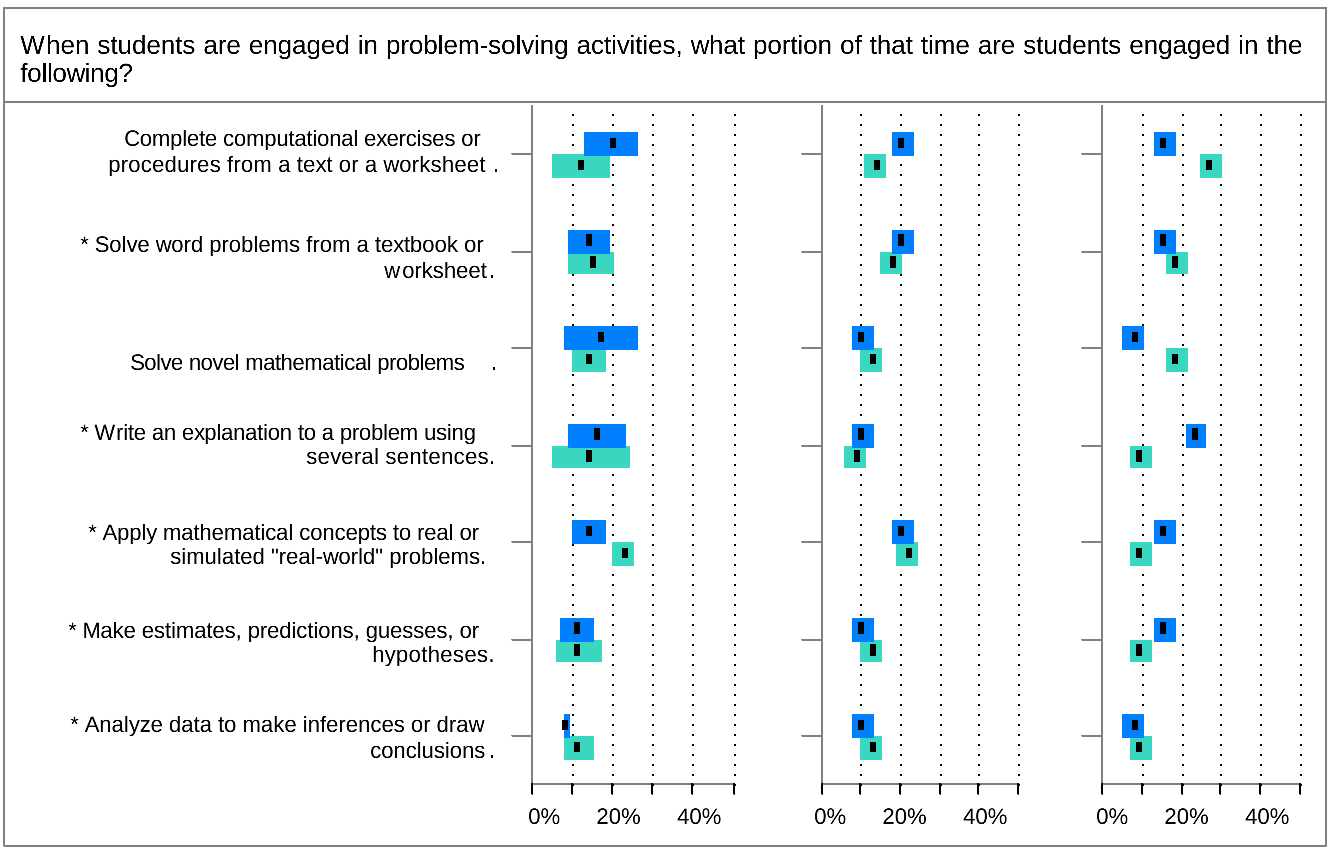
Chart 3-4
Middle School
Instructional Activities in Science
State Sub-Sample

* Item included in summary scale.

Source: CCSSO/WCER, Survey of Enacted Curriculum, 1999

Chart 3-5
Middle School Mathematics
Problem Solving Activities
State Sub-Sample



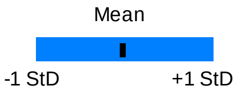


* Item included in summary scale.

Source: CCSSO/WCER, Survey of Enacted Curriculum, 1999

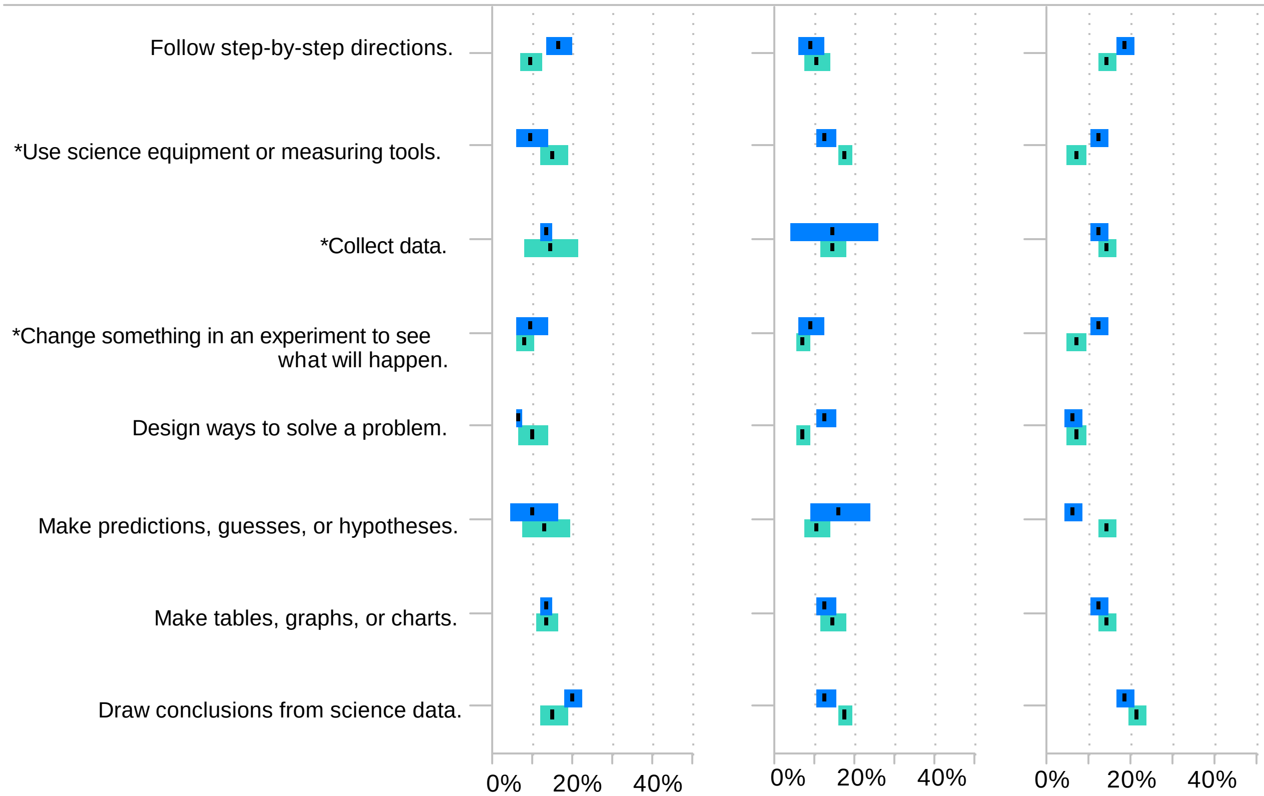
Chart 3-6
Middle School Science
Laboratory Activities
State Sub-Sample

Legend



Middle School Science		
School Comparison	**School 320**	**School 924**
School 924 School 320	Teachers C & D Teachers A & B	Teacher B Teacher A

When students are engaged in laboratory activities, investigations, or experiments, what portion of that time a re students engaged in the following?



* Item included in summary scale.
Source: CCSSO/WCER, Survey of Enacted Curriculum, 1999

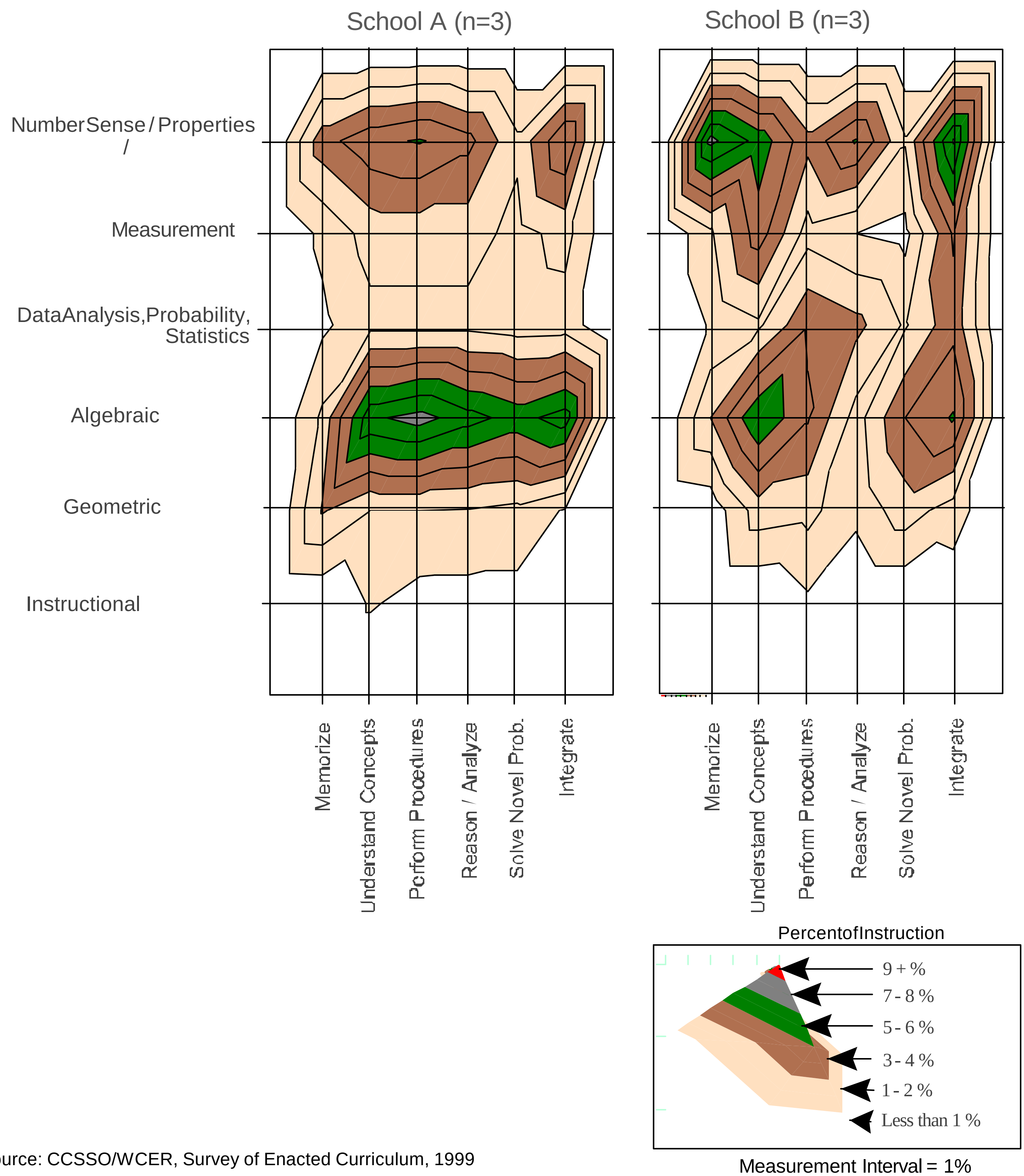
Compare Current Instruction to Standards. One of the purposes of data on classroom practices and curriculum is to address the question of alignment of instruction with state, local, or national standards for a subject. That is, how does instruction offered in our school compare to the standards for student learning under which we operate and which serve as goals for our system? The SEC data provide an excellent opportunity for educators at all levels to analyze the way that standards are being implemented in classrooms. Data reported in charts such as these included here can be used by educators to begin to analyze and understand match between standards and existing instruction.

An advantage of the content matrix section of the SEC is its consistency with standards-based learning, as described in national professional standards (NCTM, 1989; 1991; 2000; AAAS, 1993; NRC, 1995), as well as with the standards approved by most states (see CCSSO website: www.ccsso.org). The SEC content matrix is based on two dimensions: *content topic* (math or science disciplinary knowledge to be learned, e.g., geometry), and *expectations* (skills and capacities that students are expected to gain through instruction, e.g., solve novel problems). This two-dimensional approach to curriculum is common among current national and state standards.

Chart 3-7 *(Middle School Math Content Map)* illustrates the use of data from the content matrix to examine differences in content taught in classes. The percentages represented in the map are the averages for each main content topic and its intersection with each of the six types of expectations. This map shows that teachers in School A and School B spent the most time on Algebraic Concepts and Number Sense/Properties, and the teacher expectations focused on Understand Concepts, Perform Procedures, and Integrate. School A teachers placed more emphasis on Reasoning and Solving Novel Problems in Algebra, and School B teachers reported more emphasis on Memorizing and Integrating in the topic Number Sense. The maps allow teachers and administrators to quickly view the overall picture of math content and see main differences with math taught in other classrooms or schools.

Chart 3-8 (*Middle School Math Content Graph*) provides a bar graph representation of the same teacher responses as the content map. Each cell shows the average percent of time reported for that topic and expectation. The mean and standard deviation (extent of variation in teacher responses) for each topic are shown on the right Row Totals (such as, School B averaged 40 percent of time on Number Sense/Properties). The mean and standard deviation for each expectation are shown on the bottom Column Totals (such as, School A averaged 20 percent of time on Understand Concepts).

Chart 3-7
Middle School Mathematics Content Maps
State Sub-Sample

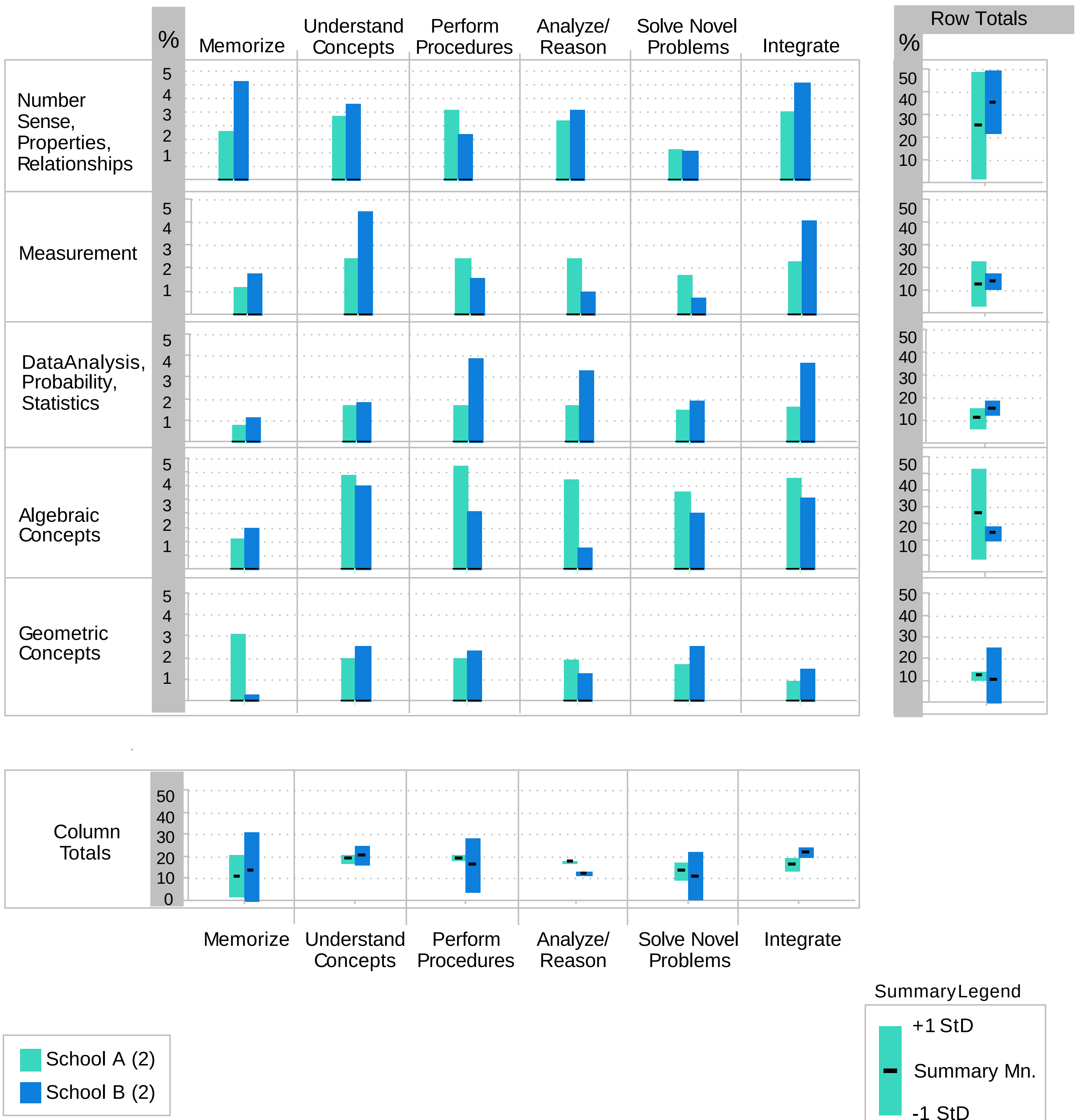


Source: CCSSO/WCER, Survey of Enacted Curriculum, 1999

Chart 3-8

Middle School Mathematics Content Graphs

State Sub-Sample



Source: CCSSO/WCER, Survey of Enacted Curriculum, 1999

Chart 3-9 *(Middle School Science Content Map)* reports the data on content of science instruction for two schools. The data are reported under six main topic categories. State and national standards typically include standards for teaching Life Science, Physical Science, and Earth Science, as well as standards for Nature of Science (including scientific method and history of science). Measurement and calculation are scientific skills included in many state and local curricula for school science.

Comparing Grade 8 science content for these two schools, we can see significant differences. Teachers in School A focused heavily on Life Science, with an average of 40 percent of time, while School B teachers reported only 20 percent of time spent on Life Science. Teachers in School B spent comparatively more time on Earth Science and Chemistry, while School A teachers spent more time on Physical Science.

Teachers also varied in the time spent on science topics. In Chart 3-10, the summary graph on the far right (Row Totals) indicates that teachers in School A varied widely in time spent on Life Science and Physical Science. School B teachers were more consistent in responses on science topics.

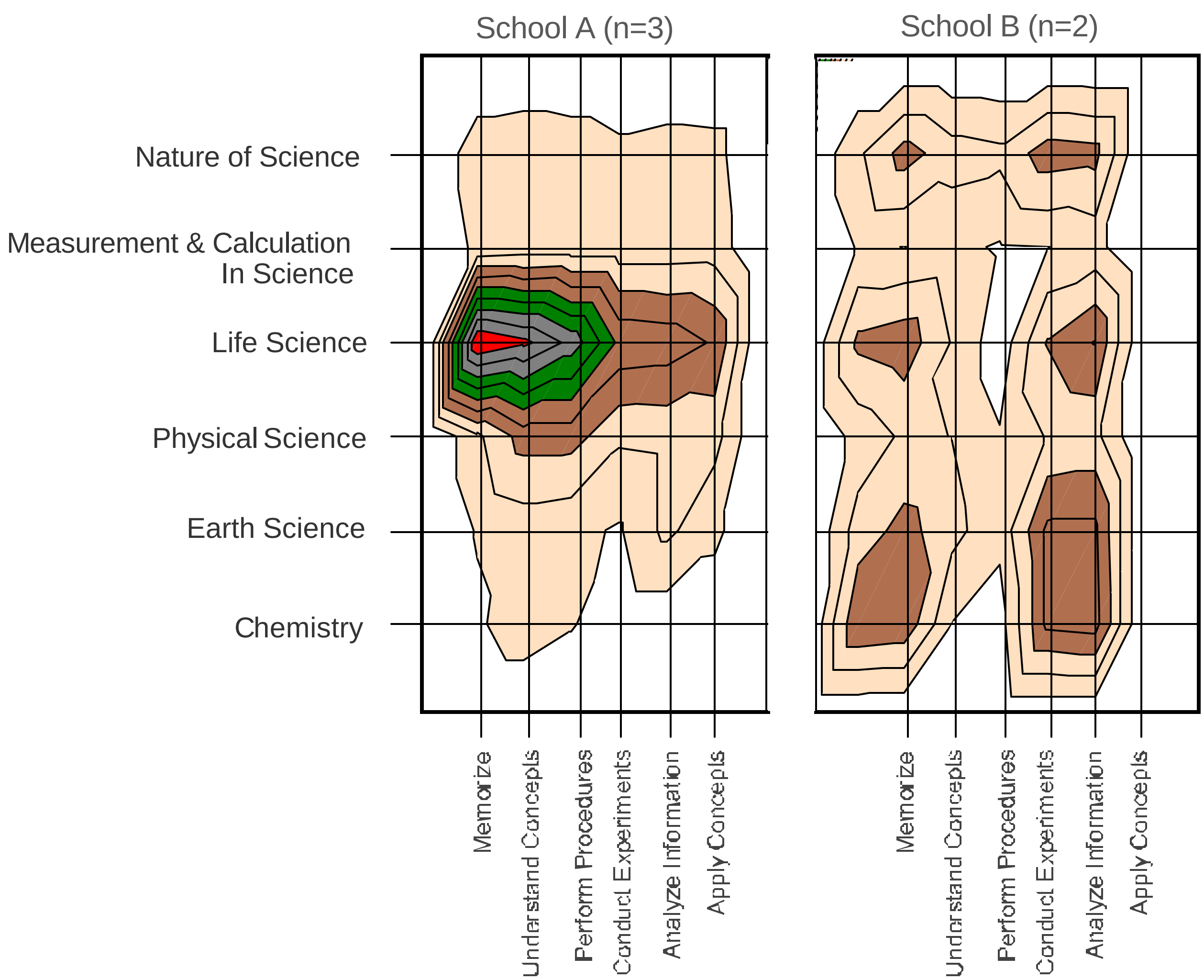
In Chart 3-10 (*Middle School Science Content Graphs*), the results show that schools differed in the *expectations for learning* in middle grades science. The data show that School A teachers reported more time spent on the expectations Memorize, Understand Concepts, and Perform Procedures than teachers in School B. These differences can be seen in the Column Totals. For example, the total time on Memorization and Perform Procedures in School A averages near 20 percent, while School B teachers reported about 15 percent on these two expectations. On the other hand, School B teachers reported more time on expectations for Analyze Information and Apply Concepts.

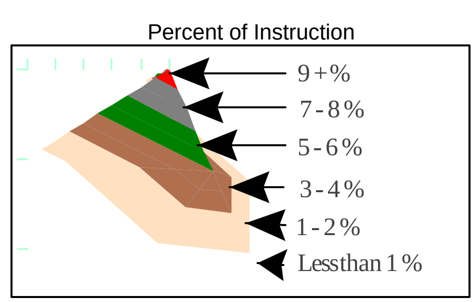
One way to analyze instructional content data is to examine the degree to which each content standard or content area is taught at a specific grade, and how the grades differ. Realistically, we would not expect the content across Life, Physical, and Earth Sciences would be taught equally each year. Ideally, the survey would be given to teachers in each grade n the same school. Then, teachers in a school or district could analyze which content /standards topics are taught and emphasized at each grade and how these vary across classes.

Chart 3-9

Middle School Science Content Maps

State Sub-Sample



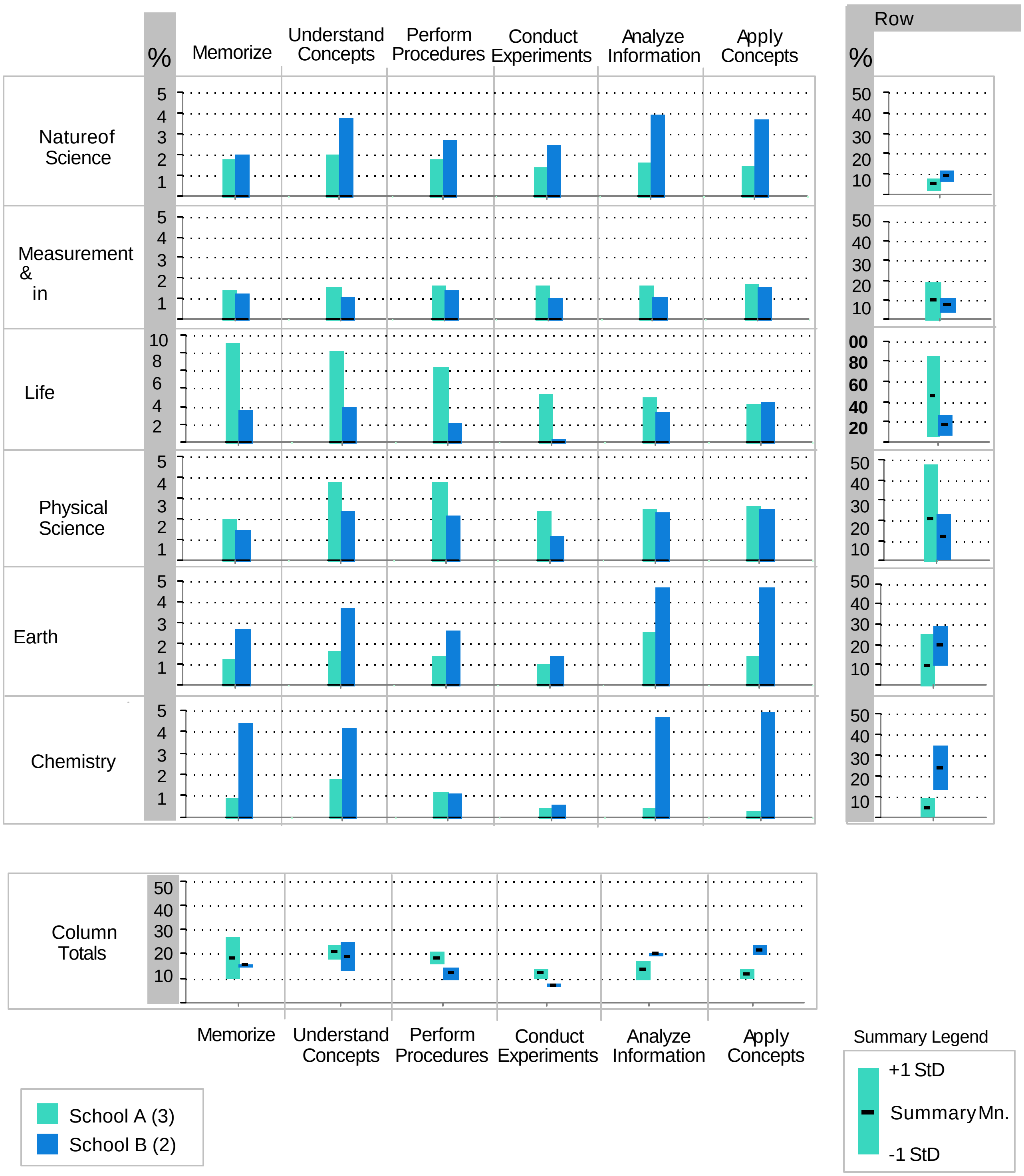


Source: CCSSO/WCER, Survey of Enacted Curriculum, 1999

Chart 3-10

Middle School Science Content Graphs

State Sub-Sample



Source: CCSSO/WCER, Survey of Enacted Curriculum, 1999

In-depth Discussion and Continued Collaboration among Teachers to Improve Practice. A further step in a school or district using enacted curriculum data is analyzing how teaching practices can be improved based on discussion and analysis of data (See Love, 2000, for thorough explanation of this model). The goal in this step is for teachers to see how their own instruction can be improved, not to make all teaching alike or to remove teacher decisions about their classrooms. The data as analyzed by teachers and others may reveal that some aspects of key district or school goals for improvement are not being implemented. Alternatively, the review of data by teachers may reveal that some teachers reported their practices incorrectly or did not understand directions (often errors in reporting are revealed only in the analysis stage). In this case, surveys may need to be repeated to obtain valid data.

After initial reviews, schools also might decide not to report data at the teacher level. The illustrations in this chapter demonstrate how individual teacher data may appear. Teachers and administrators as a group may decide that focusing on individual results is not constructive and only produces dissension and worry about poor performance evaluations. In this case, the group might decide to review the data without identifying individual teachers, thus maintaining a way to view differences in teaching approach without focusing on evaluation of performance.

As teachers review data and consider possible improvement strategies, several steps should be considered:

- Data on instructional practices can be related to student achievement, and then teachers can ask questions about the relationship of instructional methods and content to differences in achievement scores;
- Time may need to be scheduled between sessions so that teachers have an opportunity to plan together and consider their ideas and strategies for improvement;
- The data will provide initial indicators of differences in practices, leading to further information based on discussion, observation, and reflection;
- Surveys should be repeated after a predetermined period of time (e.g., one year), and then the new results should be compared with prior data to determine change in practices.

Teacher networks or work groups can be formed to continue to more closely examine these data and more detailed analyses, such as with the specific content matrix data for sub-topics, e.g., differences in what is covered in Life Science or Earth Science, or discussion of how teachers interpret expectations for student learning in terms of student performance and teaching strategies. Student work, assignments, or materials used for specific content areas or teaching activities could be shared, with teachers demonstrating how the student work relates to standards or curriculum goals.

Professional development opportunities can be planned for areas of weakness in student learning or where there is misalignment of instruction and standards. Where teachers identify areas of mutual concern about subject content that is not being learned by students, experts could be sought to address those weaknesses in the curriculum or teaching strategies.

Increasing consistency and improving articulation of curriculum is another step that can be taken by teachers and administrators based on their analysis of data. By examining the data from surveys, teachers within the same grade or course can work together to improve consistency of content. Teachers of adjoining grades and courses can use the data to plan their teaching to pick up where the curriculum left off in an earlier grade or course. Teachers can also learn to build on the content and methods of learning gained in the prior course. For example, teachers can use more complex problem solving or integration of subject areas with the knowledge that more basic, factual information or skills had been gained at the prior grades or courses. The enacted curriculum data can be an important tool to reduce the duplication and repetition of content or lower-level expectations for student learning.

Analyze Subject Content Taught in a Topic Area

The subject content portion of the SEC asks teachers to report how much instructional time they devote to topics and sub-topics within mathematics or science. The teacher responses to the content matrix section (topic by expectations for students) can be analyzed at the broad topic levels to determine overall patterns in teaching content across schools and classes, or content taught can be analyzed at the level of more specific sub-topics. For example, in elementary mathematics the main content topics are: Number Sense, Properties and Relationships; Measurement; Data Analysis, Probability and Statistics; Algebraic Concepts; and Geometric Concepts. Within each of the math topic areas, the survey content matrix requests teachers to report on from three to 17 sub-topics, such as place value, patterns, and percent under the Number Sense topic area. Then, the teachers are asked to report the amount of class time that was spent on each of the expectations for students: Memorize, Understand Concepts, Perform Procedures, Analyze/Reason, Solve Novel Problems, and Integrate.

Specific Content Taught in Elementary Math. In Chart 3-11, we show teacher-reported data at the topic level for elementary mathematics, with sub-topic data results highlighted for the Number Sense topic. We address several questions about the data: How would math educators and teachers use these data? What kinds of questions might the results from this portion of the survey answer?

Teachers reported spending the most instructional time on two topics--Number Sense/ Properties/ Relationships and Operations--and two expectations--Understanding Concepts and Performing Procedures.

- Teachers reported 23 percent of time was spent on Number Sense (which included sub-topics of Decimals, Factors, and Estimation).
- Operations was reported 22 percent of time (which included Adding, Subtracting, Multiplying and Dividing Whole Numbers, and Fractions).
- A moderate amount of time was spent on the content areas of Measurement (16%), Geometry (17%), and Data Analysis/Probability/Statistics (12%).
- Algebraic Concepts received very little instructional time (6%), and the least amount of instructional time was spent on Instructional Technology (using calculators or computers) (4%).

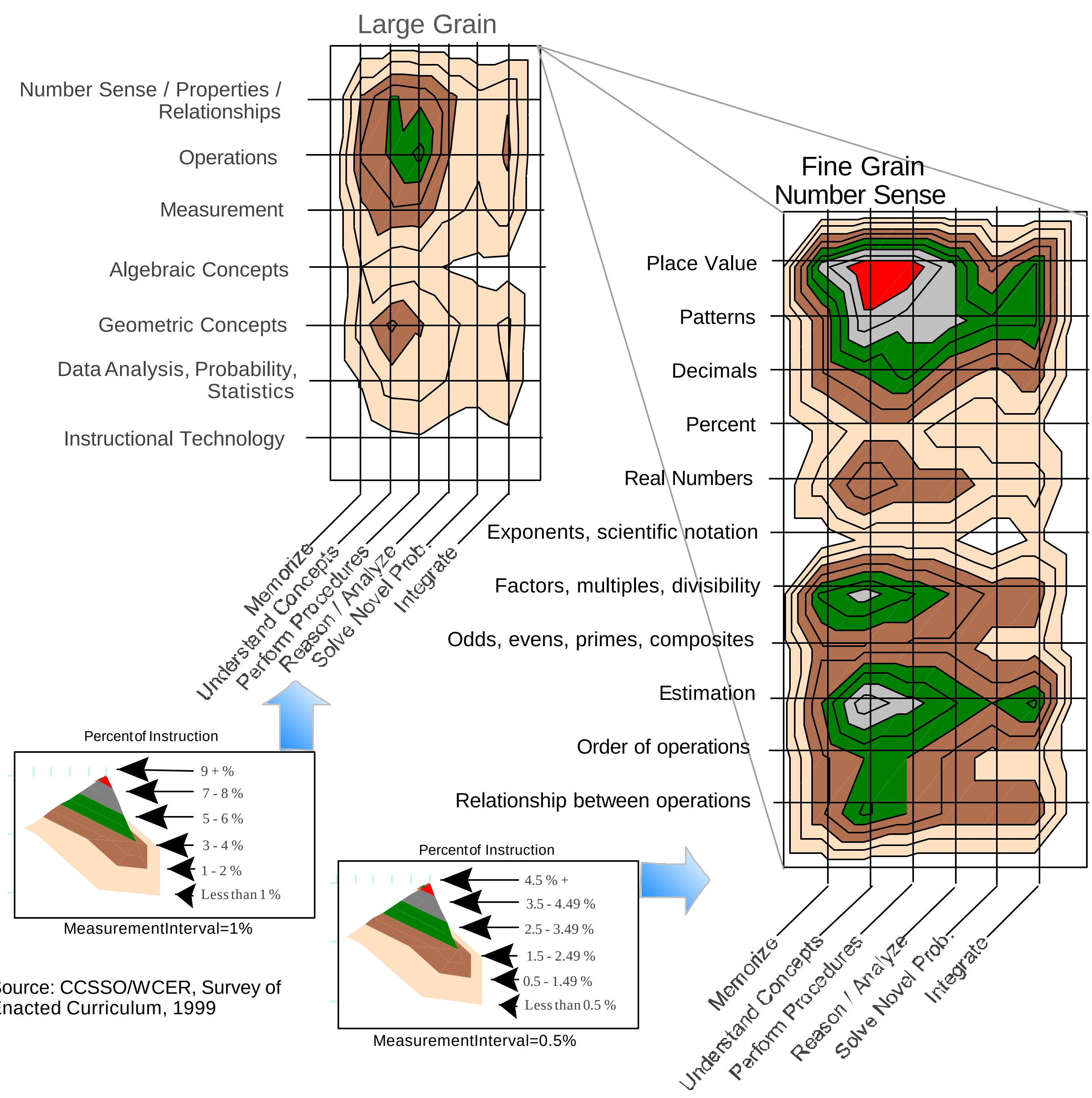
Elementary teachers reported spending 22 percent of time each on Understanding Concepts and Performing Procedures, when these were summed across topics. They spent the least time on Solving Novel Problems (11%) and Integrating (14%). In conjunction with the content emphasis, we see that teachers spent the most time on Understanding Concepts in the two content areas of Number Sense and Operations.

The results in Chart 3-11 illustrate data on time spent teaching sub-topics within Number Sense:

- Teachers spent the most time on Place value and Estimation (in Number Sense);
- Teachers reported the most time on the expectation of Performing Procedures, and little time on Solving Novel Problems or Reasoning;
- The least time is spent on teaching Decimals or Percent, and little time is spent on Fractions;
- A majority of time was spent on Operations with Whole Numbers.

These sample data are consistent with other data on math teaching at the elementary level (e.g., NAEP and TIMSS, see Wilson and Blank, 1999). Comparing the instruction data to national standards from the NCTM (2000), we found that classes spent a moderate amount of time on Measurement, Geometry, and Data/Statistics, which are three areas that typically get little attention in traditional elementary curricula. Finally, teachers spent little time with Algebraic concepts, and teachers provided little instruction where they expected students to Solve novel problems or Apply mathematics to real-world situations.

Chart 3-11
Elementary School Mathematics Content Maps
11 State Sample (n=169)



Source: CCSSO/WCER, Survey of Enacted Curriculum, 1999

Initiative vs. Comparison Schools. Chart 3-12 displays data for the Algebra topic and sub-topics grouped according to responses from teachers in Initiative and Comparison schools. Overall, the Initiative teachers spent more instructional time on the topic area of Algebra than Comparison teachers.

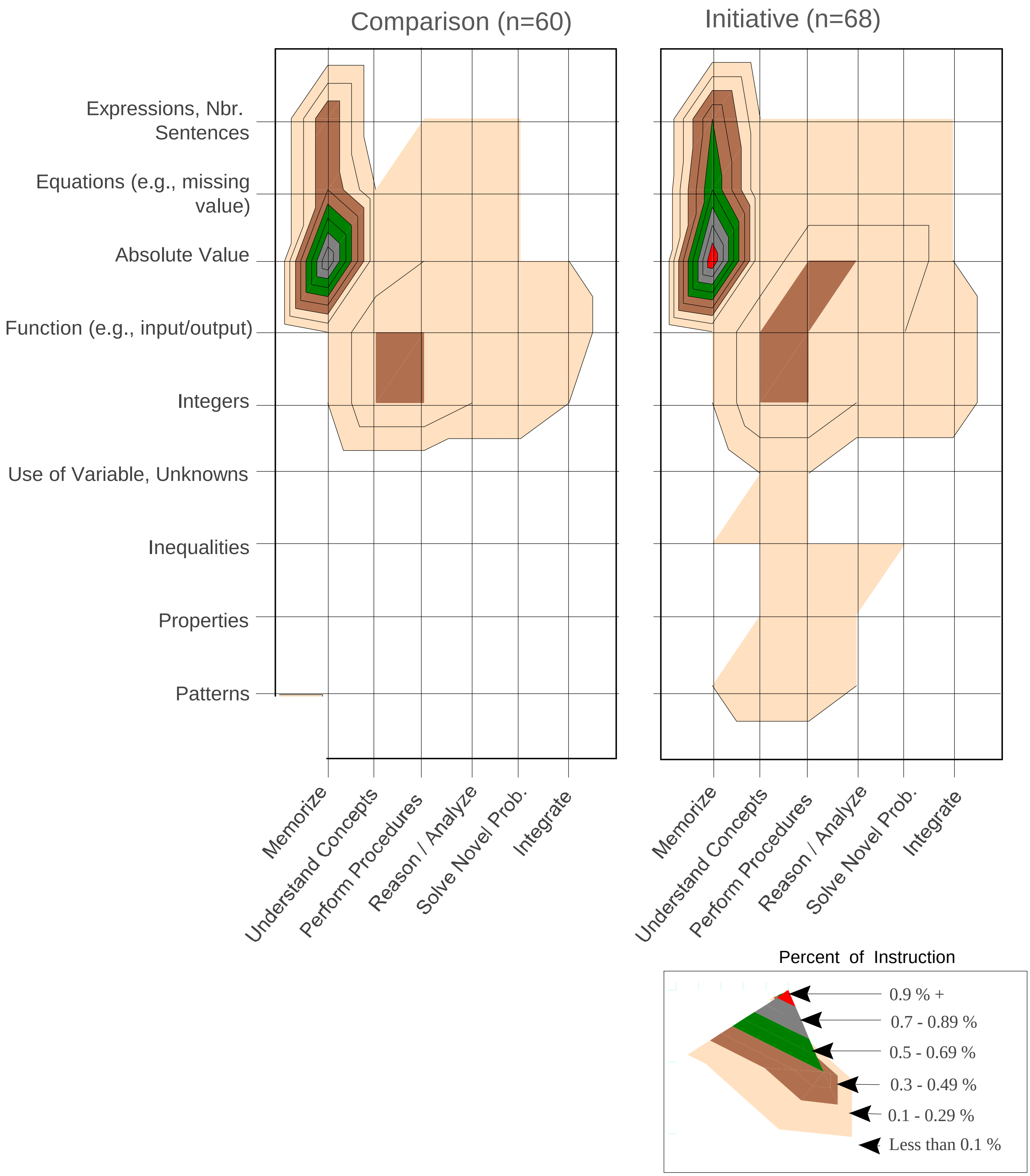
- Initiative teachers reported more time than Comparison teachers on expecting students to Analyze or Reason.
- The sub-topics within Algebra with more emphasis by Initiative teaches were Integers, Use of Variables, and Patterns.
- In teaching Integers, Initiative teachers reported more time on Understanding Concepts.
- In Patterns, Initiative teachers reported more time on Analyzing and Reasoning.

We also analyzed differences between the school groups in other math topic areas, but a chart is not displayed. In general, teachers in the Initiative schools reported more time in instruction that expected students to Reason and Analyze, across all the content topics. We found differences between Initiative and Comparison schools on instruction in the content topic of Geometry, particularly in the sub-topics of basic terms, congruence, and polygons. Initiative teachers reported more time on Performing Procedures with basic Geometric terms, work with polygons, and memorizing in the area of pie charts. In the area of Operations, the Comparison group reported significantly more instructional time on operations with equivalent fractions. In Measurement, Comparison group teachers spent more time on solving novel problems in the metric system.

The NCTM Standards (2000) and many state standards (Blank, Pechman, et al., 1997) call for the inclusion of Algebraic Concepts in the elementary curriculum. The SEC results indicate that the Initiative teachers reported more emphasis on Algebra and, likewise, reported spending more time on Geometry, another area emphasized by reform documents as important for the elementary grades. Also in line with reform expectations is the greater emphasis by Initiative teachers on Analyzing and Reasoning.

These examples illustrate the kind of analysis that is possible with data from the Survey. The data offer a unique opportunity to obtain a picture of instructional practices by examining how teaching of content and expectations interact with each other across schools and within and across different groups of teachers.

Chart 3-12
Elementary School Mathematics Content
Fine Grain Algebraic Concepts
11 State Sample



Source: CCSSO/WCER, Survey of Enacted Curriculum, 1999

Measurement Interval = 0.1%

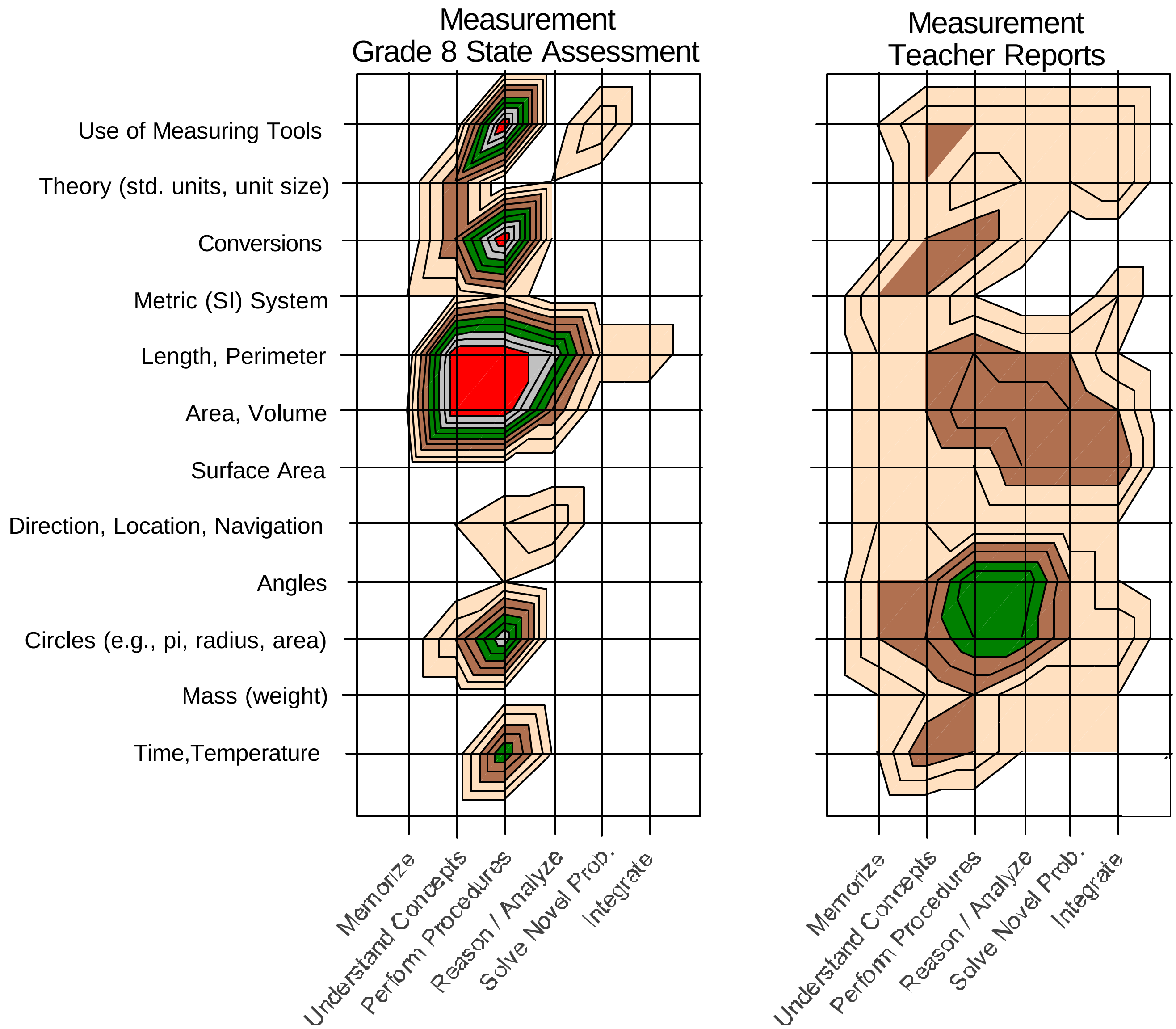
Alignment Between Instruction and Assessment

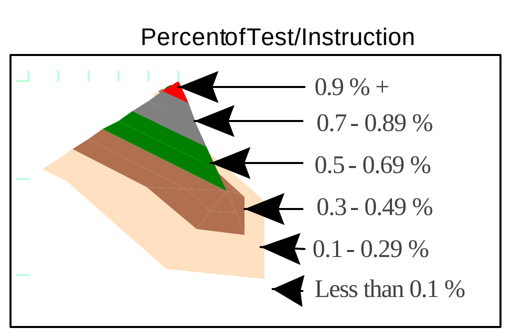
Does our state test measure the same mathematics that our teachers are teaching? The SEC data can provide valuable insights into a question such as this. One helpful tool for analyzing alignment is the content map, which can display data at the topic level (e.g., Measurement) or at the sub-topic level (e.g., Area, Perimeter). Chart 3-13 presents two content maps at the sub-topic level in Measurement at Grade 8 in Kentucky. The map at the left displays state test data in the content area of Measurement; the map at the right displays teacher-reported data on instruction in Measurement.

The first impression is that the maps are not alike. The patterns in each vary significantly. The testing map shows one primary area of concentration, while the instruction map shows several different areas with varying levels of concentration. Our first inference is that instruction is not well-aligned with instruction in this content area.

Looking more carefully at the details of the maps (horizontally), we can see that the testing map reveals that only a few content sub-topics were covered on the test, while the instruction map reveals that multiple content sub-topics were taught in the classroom. Likewise, a vertical analysis reveals that only one level of expectation predominated on the test, while instruction focused on nearly all levels of expectation. In particular, the test items concentrated on Length and Perimeter, Area and Volume, and on Performing Procedures as the expectation for students. We can imagine that such items require students to calculate these measurements by following routine procedures or formulas. In contrast, instruction seems to have addressed nearly all the measurement sub-topics in the survey, with the exception of Mass. Nearly all levels of expectation, from Memorizing to Solving Non-Routine Problems and Integrating Measurement with other topics, are covered to an equal degree.
Our findings from these sample data are consistent with other evidence about mathematics instruction, such as from TIMSS results (NCES, 1996, 1997). The data show that instruction is spread widely over many topics, but the depth of teaching expressed as expectations for students is not deep. Data on the sample state test are consistent with common complaints about large-scale assessments by its heavy focus on students performing mathematical procedures.

Chart 3-13
Grade 8 Mathematics
Measurement Sub-Topics
Alignment Analysis
State Sub-Sample





Source: CCSSO/WCER, Survey of Enacted Curriculum, 1999

Implications of Alignment Analysis. First, state leaders might want to ask themselves if their goal should be to have perfect alignment between instruction and a state-level test. While alignment is generally seen as a good thing, one should be cautious in seeking to align instruction too rigidly with state-level assessment. After all, a state test is typically bound by certain parameters, such as being a timed, paper-and-pencil exam. Such tests cannot adequately assess all of the important sub-topics in mathematics. More importantly, there are expectations that are very difficult to assess in any kind of test. Analyzing, reasoning, solving non-routine problems, and integration can be performed to a limited degree on a written exam, but certainly not every sub-topic that is important to teach can be tested this way. Determining whether a student is able to choose appropriate measuring instruments and then use them correctly might be done more effectively by the classroom teacher who can observe the student. And expecting teachers to teach only what can easily be assessed on a large-scale test would result in a harmful narrowing of the curriculum.

Nevertheless, an alignment analysis such as this one can point to gaps on either the test or the curriculum. Perhaps the test items need to be broader in scope or assess more than one expectation. Or perhaps instruction needs to be more focused on certain topics that are deemed more important than others. Perhaps more instructional time needs to be spent overall in a certain topic area. Alignment Analysis offers the two-fold benefit of providing comparisons of current classroom practices in relation to policy documents such as standards and assessments, as well as allowing teachers to compare their own instructional emphases to those of other teachers.

Chapter 4: Quality of Data

The results reported in Chapters 2 and 3 hinge upon the quality of data collected through surveys. Since the SEC instruments depend upon teacher self-report, the reliability and validity of data from the surveys need to be addressed to ensure confidence in the resulting descriptions and analyses. The sampling plan and data collection procedures for the study are key factors in establishing the extent to which the results can be generalized beyond the sample.

In this chapter, characteristics of the data set used in the Study of Enacted Curriculum with the participating 11 states are discussed, and key issues affecting the quality of teacher-reported data on the enacted curricula are raised. Suggestions are provided for improving data quality using the SEC instruments.

Using Teacher Reports of Practice

Prior to the use of content standards as policy tools for implementing system-wide reforms in education, education policy research treated classroom instruction as a "black box" that was not susceptible to wide-scale analysis. With the advent of standards and a push toward more challenging content for all students, studying differences in classroom practices and instructional content have become central to research, and the "black box" can no longer remain unexamined. Yet direct observational study of classroom practices is no small undertaking.

To capture the rich and complex dynamics of the classroom environment typically requires a team of researchers that is only able to study a handful of classrooms over extended periods of time, with use of a case-study methodology. Such in-depth examinations are of course impossible on any large scale, and it is large-scale descriptions that are necessary for evaluating the effectiveness of reform programs. The logical alternative is to utilize an indicator system based on a finite set of descriptors that can be reported on by teachers in a survey format. The use of teacher self-report data collection raises important questions about teacher candor and recall, as well as the adequacy of the instrumentation to provide useful descriptions and indicator measures.

Teacher Candor. Some educators and researchers are concerned that teacher reports may be biased toward "socially acceptable" responses. Certainly if questions of practice are asked in a high-stakes environment, where the answers given might be expected to fulfill some accountability procedure, there would be cause for concern. If teachers believe their response to a survey might impact their livelihoods in some way, there is good reason to worry about the candor teachers will exercise in reporting their practices. For this study, teacher reports were collected on a voluntary basis, and the teachers were guaranteed anonymity and were clearly told that the data would remain unrelated to any accountability system or policy.

The best means for ensuring teacher candor is to make the exercise of completing the survey of personal value to the individual teacher by making it possible for teachers to gain confidential access to their own results for personal reflection. At a minimum, teachers should be provided concise reports for their school, district, and state. We recommend in addition that teachers be provided training on using these types of data as tools for personal and school level improvement (see Chapter 3).

***Reporting Accuracy*.** Even with a teacher's best effort to provide accurate descriptions of practice, those descriptions are constrained by the teacher's ability to recall instructional practice over extended periods of time. Daily teacher logs accumulated over a school year provide the best source for detailed descriptions of practice, but these are expensive and burdensome. Survey reports covering a semester or school-year are more economical, and less of a burden on teachers, but do require that teachers be able to recall aspects of instructional practice months after many of those activities had occurred.

Teachers participating in the 1999 Study reported on a full school-year of teaching in science or math, as has been done with previous administrations of the instruments. The best validation of this approach, requiring teachers to recall instruction for a full school-year, comes from the Reform-Up-Close study (Porter, et al., 1993) where researchers collected and compared daily logs, independent observation, and teacher survey reports (Smithson and Porter, 1994). The study found that data reported about curriculum content in teacher surveys covering a whole year were highly correlated with the data from daily logs of instructional content.

However, teacher surveys not only require accurate recall by teachers but also common understanding of what is meant by key terms describing instruction. Teachers may think they are doing one thing, when an independent observer would characterize their activities differently (Cohen, 1990). From his observational studies, Knapp noted “they (teachers) know the words but they can’t sing the tune of standards-based practice (1996).”

In using survey instruments such as the SEC over the past fifteen years, there has been an increase in the tendency for teachers to report a balanced curriculum across categories of student expectations or cognitive demand, that is, teachers report some expectations for students in all of the categories (Smithson and Porter, 1994). The results from the 1999 data also show this pattern of teachers reporting all categories. While we do not question teacher candor, we do wonder about their accuracy. It is possible that discussion by teachers of the meaning of the various categories of cognitive demand and the content consistent with those categories, coupled with teachers observing fellow teachers, could provide a very powerful tool for professional development.

Adequacy of the Language. The language of description is a crucial element in any indicator system. For theory-driven research, the language of description is typically tailored to fit the theoretical constructs—subjects and objects relevant to the theories/hypotheses being investigated. The SEC instruments by contrast are intended to be appropriate for a broad range of instructional practices. The language utilized provides descriptive measures that are intended to be comprehensive, allowing teachers to find familiar ways of describing their practices, whether using traditional or reform-oriented instructional approaches. The SEC instruments have undergone numerous field tests and piloting by teachers, and they have been revised and improved at each stage. The language employed by the SEC continues to be examined and refined in order to ensure that the terms are interpreted similarly across teachers.

Sampling

Selection criteria, response rates, and sample size are important in considering the extent to which particular data can be generalized. Although a state-representative sample would be most desirable for evaluating statewide initiatives or standards, such large-scale sampling would require significant state resources. Fortunately, the SEC data can provide useful information at a smaller scale to inform both personal enrichment and school improvement decisions. The data can be powerful whether reported at the level of district, school, or individual teacher.

Selection criteria. State leaders were asked to select schools and teachers to participate in the study based on a state math and/or science policy initiative (see Appendix B). The initiatives varied by state, but methods of sampling were consistent. Each state was asked to include schools from urban, suburban, and rural districts, and schools were to vary in size and student composition. All but two states selected and assigned schools based on two categories: comparison and initiative, with schools from each category matched on enrollment, poverty level, and geographic location. Ohio and Pennsylvania chose not to differentiate between initiative and comparison schools, as their selected initiatives were intended to reach all schools in the state.

For the participating states, this selection procedure was acceptable and preferable, since each state was able to examine differences in practices in relation to their own state initiative. However, for the purpose of conducting a multi-state study, these selection criteria were less than ideal, because the criterion for assignment to the comparison or initiative group varied by state. As a result of the problems in comparability, the CCSSO/WCER research staff decided to use the number of hours of professional development reported by teachers as a standard measure for grouping teachers into the initiative and comparison categories.

Response Rate. The initial study design called for 20 schools per subject, at each of two grade levels (grades 4 & 8), for a total of up to 80 schools per state (40 mathematics and 40 science). State education staff were responsible for identifying schools that participated in programs linked to the state initiative, and they were also responsible for recruiting schools to conduct the surveys. Some state staff sent letters to districts and schools inviting participation but making it a voluntary activity. Other state staff did relatively little to recruit schools and raise interest in the project.

In general, the 11 states did not meet the target numbers in the study design. Table 4-1 summarizes the number of schools and teachers that were sent surveys, and the number of completed surveys received in return. A total of 471 schools and 1251 teachers were in the study sample. Responses were received from a total of 626 math and science teachers. Note that some of the states had small numbers of schools, and states typically selected the same schools for math and science surveys. One of the 11 states (North Carolina) focused on studying a mathematics initiative and did not collect data from science teachers.

Across the states, a total of 288 science teachers responded to the survey (46% of the 623 that were sent, and 36% of the initial target of 800 science teachers). In mathematics, 338 surveys were completed (54% of 630 sent, or 42% of the target). At least two teachers per school were invited to participate for each subject area (mathematics or science). Some schools chose to invite more teachers to participate, and some schools chose to include both mathematics and science teachers. Teachers from schools designated as Ainitiative@ were slightly more likely to respond (46% of total responses) than schools designated as comparison (35%). (Note: schools and teachers were not aware of their category—initiative vs. comparison.) The remaining 19 percent of respondents came from the two states that did not identify initiative or comparison schools. Further descriptive characteristics of the sample of teachers and schools participating in the study are provided in Appendix A.

Table 4-1
Response to Survey of Enacted Curriculum by State, Spring 1999

State	Number of Sample Schools	Teacher Surveys Sent	Completed Teacher Surveys	Math Completed	Science Completed
Iowa	57	112	75	55	20
Kentucky	65	145	54	24	30
Louisiana	28	103	46	20	26
Massachusetts	40	155	95	51	44
Minnesota	37	76	43	21	22
Missouri	33	115	40	23	17
North Carolina	25	45	42	42	*
Ohio	41	125	56	29	27
Pennsylvania	28	103	46	20	26
South Carolina	40	106	52	*	52
West Virginia	77	166	77	53	24
Total	**471**	**1251**	**626**	**338**	**288**
Response Rate			50%	54%	46%

* NC did not survey science teachers;
SC conducted a math teacher survey in Fall 1999.

These response rates were well below desirable levels. There are a number of reasons for the low response rate. First and foremost, survey administration was de-centralized, handled individually by each state, and no formal follow-up procedures were instituted to track teacher completion in order to improve the response rate. Also, while some schools agreed to participate, this did not mean that their teachers had made the same agreement. In most cases, teachers were handed an envelope containing the survey with a cover letter explaining the voluntary and confidential nature of teacher participation. As a result, in some cases schools that had volunteered to participate ended up having no teachers return surveys. Moreover, the SEC instruments can take from 60 to 90 minutes to complete, and many teachers were not willing to put the required amount of effort into completing and returning the surveys. (See the section below on administering surveys for suggestions on improving teacher response rates using these instruments.)

Sample Representation. While the sample for this study was not intended to be representative of any of the participating states, alignment analyses of instruction across states and within states do suggest that the results would not be different if representative samples had been collected. When survey data on instructional content are aggregated by state, and then compared to one another using the alignment procedures described in Chapter 2, the levels of between-state instructional alignment are surprisingly high (see Table 4-2). For elementary math, the average state-to-state alignment of instructional practice is .80. For elementary science the average between-state alignment was .70, and for Grade 8 mathematics it was .68. The lowest between-state alignment of instruction was found in Grade 8 science, with an average alignment of .64. This is most likely a result of the broad number of potential topics that can be taught as part of Grade 8 science. When individual teacher reports rather than state averages are compared, the alignment measures drop dramatically (e.g., alignment index among Grade 4 math teachers = .49). This pattern of higher alignment when teacher reports are aggregated by state and lower alignment among individual teachers suggests that the sampling procedures we used did yield fairly stable results.

Table 4-2
Alignment of Instructional Content based on Teacher Survey Results

Grade/Subject	**State to State**	**Teacher to Teacher**
Grade 4 Mathematics	0.80	0.49
Grade 8 Mathematics	0.68	0.47
Grade 4 Science	0.70	0.38
Grade 8 Science	0.64	0.36

Validity

In order to judge the validity of survey-based, self-report data, it is useful to have other sources of information to draw upon in order to compare results across different data sources. Classroom observations are typically chosen as a method of confirming information on classroom practices. Unfortunately, classroom observations are a poor source for validating year-long survey reports, particularly reports of instructional content, as observations as a practical matter can only cover a very small portion of the curriculum covered over the course of a school year.

Teacher logs provide another potential source for validating information on survey reports, but these still require regular, year-long reporting by teachers throughout the year in order to accumulate log data across the school year. Teacher logs are also based on teacher self-reports, and so the question of teacher candor could still be a consideration in interpreting the results. Nonetheless, log data do allow for more frequent reporting by teachers, and thus permit useful comparisons between frequent (e.g., daily or weekly logs) and infrequent (e.g., semester-long or year-long survey reports) data collection instruments. Such comparisons can help to determine if teacher recall is adequate for using year-long survey reports. Comparative analyses between observations and daily logs, and between aggregated daily logs and survey reports, were conducted on early versions of survey-based indicator instruments as part of the *Reform-Up-Close* study, with good results (Porter, 1993).

Students provided yet another source for collecting information on classroom practice. While student reports may differ from teacher reports insofar as students may perceive instruction or interpret survey questions differently from teachers, they do provide a useful source for comparative, and potentially validating information.

Student Surveys. Student data were collected from 123 classrooms (60 science and 63 mathematics) of teachers participating in the 11-state study. Correlations were computed between student and teacher responses in order to determine degree of consistency between student and teacher reports. Student data were aggregated by class, so that comparisons could be made between the teacher reports and the class average from student reports. Results of these analyses differed dramatically by subject area. Among the mathematics classes, student and teacher reports correlated well. Indeed, of the 49 survey items for which student and teacher items were comparable, all but three items had significant and positive correlations (ranging from .20 to .74). For science, student and teacher reports significantly correlated for only 28 of the 49 survey items. The strongest levels of agreement between student and teacher reports in science were on items related to activities associated with group work, collecting information, and computer use. The lowest levels of student and teacher agreement in science were in connection with estimating the frequency of laboratory activities, portfolio use, teacher demonstrations, and writing assignments in science.

While student reports provide supporting evidence for reports by mathematics teachers, it is not immediately clear how to interpret the results for science. It may be that mathematics classes are more structured, with only a few common classroom activities that are repeated throughout the school-year (e.g., computation and problem-solving), while science related activities may be less consistently organized, requiring more imprecise estimations from both teachers and students. Or it may be that the language used for describing science instruction is less familiar to students, or students may simply interpret key terms in different ways than teachers. For example, portfolios may mean something quite different to students than to teachers. Another possibility is that in science not all students in a class have the same instructional experience. In order to test any of these explanations, other types of data would need to be collected (e.g., classroom observations and/or student and teacher interviews). This demonstration study did not attempt to collect these or other qualitative measures that would help in interpreting and validating the survey results.

Reliability and Validity for Content Analyses of Tests. In order to calculate measures of alignment between instruction and assessments, content analyses were conducted with state assessments and NAEP mathematics and science assessments. To conduct these analyses, subject matter specialists were brought together for a two-day workshop. The raters were trained in the use of the content taxonomies and on the procedures for coding assessment items. At least four raters did independent coding for each test analyzed. The results among raters were then compared using similar procedures as developed for measuring alignment, and the resulting statistic established the level of inter-rater agreement in content analyzing the various assessment instruments.

In interpreting these inter-rater agreements, it is important to realize that any one item could potentially assess several different types of content. The procedures limited raters to selecting only three topic-by-cognitive demand combinations per item. This undoubtedly forced some disagreements among raters. When making distinctions at the finest grain of topic, alignment was in the neighborhood .40 to .50. Since assessments were described as the average across raters, and each form was content analyzed by at least four experts, the reliability of the descriptions of the tests are considered to be high.

Administering Survey of Enacted Curriculum

While the response rates for teachers invited to participate in the 11 state Study were lower than desirable, the Study did provide a number of useful lessons in improving the response rate among teachers. For example, the poorest response rates were seen in those schools where teachers were given the surveys to complete on their own, at their convenience. The best response rates came from those schools in which teachers were gathered as a group for the express purpose of completing the instruments. Response rates were also higher in those districts where teachers were compensated or given professional development credit for the time it took them to complete the survey.

Teachers are also more likely to take the time and make the effort to complete the survey when they perceive some personal value to the information they provide. For this reason it is recommended that teachers be provided with confidential, individual reports that allow teachers the opportunity to reflect upon their descriptions of practice as revealed by our reporting formats. It is also strongly recommended that data on the enacted curriculum not be included as a part of any accountability system, as this could dramatically affect teacher candor in responding to the survey questions. Finally, consideration should be given to providing teachers with the results of content analyses of their high stakes tests (where used) in order to better target their coverage of assessed content and increase the value of participation for individual teachers.

APPENDIX A

Descriptive Data on Schools, Teachers, and Classes Participating in the 1999 Survey of Enacted Curriculum

Number of States	**11**

Teachers Responding to Survey
(whole or in part)

Teachers	Mathematics	Science
Elementary	169	156
Middle	137	115
High	32	17
Total	338	288

Schools by Enrollment
(selected as Initiative vs. Comparison)

	Schools with Math or Science Surveys		
	Initiative	**Comparison**	**Not reported**
Under 300	11	10	3
301 - 600	49	38	20
601 - 1000	42	38	7
1001 or more	19	8	2
Not reported	28	22	30
Total	149	116	62

School % Poverty

	Schools with Math or Science Surveys		
	Initiative	**Comparison**	**Not reported**
0 - 34%	77	46	13
35 - 74%	62	64	12
75 - 100%	8	1	7
Not reported	2	5	30
Total	149	116	62

Class Reported by Teacher
(1 class per teacher)

	Math	Science
Grade 3	16	14
4	155	126
5 or 6	4	10
7	19	20
8	111	75
9 or higher	30	10

Teaching Time
Elementary (hours/week)

	Math %	Science %
Less than 4	4.8	44.9
4 - 4.9	16.2	31.5
5 or more	79.1	23.5

Teaching Time (continued)			
Middle Grades (hours/week)		**Math %**	**Science %**
	Less than 4	20.2	13.4
	4 - 4.9	31	34
	5 or more	49	52.6

Achievement Level of Students			
		Math %	**Science %**
	High	16	9.7
	Average	50.2	47.2
	Low	13	14.1
	Mixed	19.6	29

Teacher Characteristics			
Experience: Yrs. in Subject		**Math %**	**Science %**
	0-2	12.8	11.8
	3-5	18.6	13.7
	6-11	21	22.6
	12 or more	47.6	51.9
Major: Bachelors or Highest		**Math %**	**Science %**
	Elementary Ed.	40.6	43.5
	Middle Ed.	6.7	4.2
	Math Ed.	13.3	
	Science Ed.		12.6
	Mathematics	10.5	
	Science field		11
	Other	28.9	26.7
		Math %	**Science %**
	BA/BS	51.4	42.4
	MA/MS or higher	48.7	57.6

Teacher Professional Development			
Content study in field (hrs. in last year)		**Math %**	**Science %**
	< 6	32.6	52.8
	6 - 15	25.2	21.6
	16 or more	22.2	25.7
Methods of teaching in field		**Math %**	**Science %**
(hrs. in last year)	< 6	47	34.9
	6 - 15	29.4	46.7
	16 or more	23.7	18.4

Teacher Demographics			
		Math %	**Science %**
	Female	82.1	76.8
	Male	17.9	23.2
	White	93.9	90.7
	Minority	6.1	9.4

APPENDIX B
Information on State Initiatives, Standards, and Assessments

The study of state reform is based on a design for surveying a selected sample of teachers and analyzing the data to determine effects of a state initiative in mathematics or science education on teaching practices and curriculum. In some states, the initiative is directly linked to state content standards. In others, the initiative relates to a broader set of state education policies to improve education. Six of the 11 states were in the state systemic initiative (SSI) program supported by the National Science Foundation. The following is a summary outline of key state information upon which the study is based.

Iowa		
Surveys, Spring 1999	*Grades*	
Mathematics	4, 8	
Science	4, 8	
State Initiative		*Year Implemented*
Mathematics:		
First Governor's Conference on Reform in Math Ed. (K-12)		1992
Science:		
New Standards project		1992
Science SCASS Assessment project		1993
Nat. Science Ed./NCTM Standards Awareness / Implementation		1996
State Content Standards		
(Standards and Frameworks developed at the District level)		
State Assessments		
(No statewide assessment)		

Kentucky

Surveys, Spring 1999	*Grades*	
Mathematics	4, 8	
Science	4, 8	

State Initiative	*Year Implemented*
Appalachian Rural Systemic Initiative (ARSI)	
KERA -- State Reform	1990
Partnerships for Reform Initiatives in Science and Mathematics (PRISM)–NSF/SSI	1991
Kentucky Middle Grades Mathematics Teacher Network	4 years
Eisenhower Regional Consortium for Mathematics and Science Education at AEL	1993
Informal Science Organization/School Partnerships	
K - 4 Mathematics Specialist Program	3 years

State Content Standards	
Transformations: KY Curriculum Framework	1995
KY Core Content for Math and Science Assessment	1996
Program of Studies for KY Schools, PK-12	1997

State Assessments	
KIRIS Math/Science Gr. 4, 8	1991
Commonwealth Accountability Testing System (CATS)	
Math Gr. 5, 8; Science Gr. 4, 7	1998

Louisiana

Surveys, Spring 1999	*Grades*
Mathematics	4, 8
Science	4, 8

State Initiative	*Year Implemented*
Math & Science: LA Systemic Initiatives Program	1991
K-3 Reading and Mathematics Initiative	
Developing Education Excellence and Proficiency	

State Content Standards	
LA Mathematics and Science Content Standards	1997
LA Mathematics and Science Curriculum Frameworks	1995

State Assessments	
Math: CRT Gr. 4, 8	NRT Gr. 3, 5, 6, 7, 9
Science: CRT Gr. 4, 8	NRT Gr. 3, 5, 6, 7, 9

Massachusetts		
Surveys, Spring 1999	*Grades*	
Mathematics	4, 8	
Science	4, 8	
State Initiative		*Year Implemented*
Partnerships Advancing the Learning of Math and Science (PALMS):		1992
PK-12/ Higher Education Goals: Increase student achievement in math, science and technology. Reduce achievement gaps for ethnic, bilingual, and gender groups		
Focal Areas:		
1) Data-driven systems		
2) Standards-based curriculum, instruction, assessment		
3) Qualified/quality teachers		
4) Middle school/high school and transitions		
5) Parent/community involvement		
State Content Standards		
Mathematics Curriculum Framework: Achieving Mathematical Power		1995
Science and Technology Curriculum Framework: Owning the Questions through Science and Technology Education		1995
State Assessments		
MCAS: Math, Science: Grades 4, 8, 10		1997-1998

Minnesota		
Surveys, Spring 1999	*Grades*	
Mathematics	4, 8	
Science	4, 8	
State Initiative		*Year Implemented*
Minnesota Graduation Standards		Fall 1998
Basic Standards (R, W, Math) and High Standards in 10 areas including Math and Science.		
Math reform schools: Have implemented NSF math curricula. Non-reform: traditional, textbook programs.		
Science reform schools: Some, not all using kit-based programs--FOSS or STC; Others, with teachers in the "best practice network in science."		
State Content Standards		
Mathematics K-12 Curriculum Framework		1997
Science K-12 Curriculum Framework		1997
State Assessments		
Basic Stands. Test	Math, Gr. 8	1996
Comp. Assess.	Gr. 3,5	1997-98

Missouri

Surveys, Spring 1999	*Grades*
Mathematics	4, 8
Science	3, 7

State Initiative

"Initiative" schools were selected from two programs:

(1) Schools that participated voluntarily in the first year that state-level performance-based math and science assessments became available. Mathematics assessment began Spring 1997, and implemented statewide 1998. Science assessment began Spring 1998.

(2) Schools in districts that participated in the voluntary inservice training program on performance-based assessment, which began being offered by the state in 1993.

State Content Standards	*Year Implemented*
Math Curriculum Frameworks	1996
Science Curriculum Frameworks	1996
Show Me Standards	1996

State Assessments

Missouri Assessment Program	
Math: Grades 4, 8, 10	1996-97
Science: Gr. 3, 7	1997-98

North Carolina

Surveys, Spring 1999	*Grades*
Mathematics	4, 8

State Initiative

"The ABC's of Public Education" is an initiative that began in 1992. There are three parts: A: Accountability; B: Basics and High Education Standards; and, C: Maximum Local Control. Key aspects of each part:

> A -- Individual schools held accountable, staff responsible, students tested in grades 3-8, high school end of course tests, and schools judged on raising achievement.
>
> B -- State standards in Reading, Writing, and Mathematics; and grade specific objectives per content area and tests based on objectives.
>
> C -- Principals and teachers make decisions on materials and instruction; state provides information on "best practices," curriculum standards, and technology.

State Content Standards	*Year Implemented*
NC Standard Course of Study, Mathematics Competency-based	
Curriculum Teacher Handbook K-12	1994
Strategies for Instruction in Mathematics (state-provided for each grade)	

State Assessments

North Carolina Testing Program: Math: Gr. 3-8, Algebra	1992

Ohio

Surveys, Spring 1999	*Grades*
Mathematics	4, 8
Science	4, 8

State Initiative	*Year Implemented*
Urban Schools Initiative (USI)	1996

The USI was launched by the Ohio Department of Education to comprehensively address the challenges facing urban school communities. The initiative represents all twenty-one of Ohio's urban school districts, 24% of the state's total student population, and 72% of its minority students. The USI has been a leader in developing and implementing new programs, attracting grants, and making a positive impact on students. With its District Team, School Readiness Resource Centers, Professional Development and Disciplinary Intervention Grants, and its widely circulated report, *Through the Eyes of Children*, Ohio's USI has had a substantial impact on the state's urban school communities.

State Content Standards

Model Competency-Based Mathematics Program	1990
Model Competency-Based Science Program	1994

State Assessments

Ohio Proficiency Test Program

Mathematics: 4, 6, 9, 12

Science: 4, 6, 9, 12

Pennsylvania

Surveys, Spring 1999	*Grades*
Mathematics	4, 8
Science	4, 8

State Initiative	*Year Implemented*
Not reported	

State Content Standards

Mathematics Curriculum Framework	1996
Science and Technology Framework	1995
Mathematics Standards	1999

State Assessments

Math: NRT Gr. 5, 8, 11

Science: Developing

South Carolina

Survey	*Grades*	
Science, Spring 1999	4, 8	
Mathematics, Fall 1999	4, 8	
State Initiative		*Year Implemented*
SC State Systemic Initiative (K-12) -- reform, standards implementation		1992
Teachers Leading Teachers (Gr. 4-8) -- reform, physical science content		
Science SCASS Assessment project (K-12)		1995
Nat. Sci. Standards-- Building a Resource (K-12)		
Instructional Improvement Initiative (K-12) -- Low performing schools, to increase student achievement		
State Content Standards		
SC Science Framework		1996
SC Acadademic. Achievement Standards for Science		1996
SC Science Curriculum Standards		1998
Expert review and Revisions to Science Standards		1999
Mathematics Standards		1998
Mathematics Framework		1993
State Assessments		
Basic Skills Assess.		
Math: Gr. 3, 8, 10		
Science: Gr. 3, 6, 8		

West Virginia

Surveys, Spring 1999	*Grades*	
Mathematics	8, 10	
Science	8, 10	

<u>*State Initiative*</u>	<u>*Year Implemented*</u>
Mathematics Reform: Project MERIT	1998
(Mathematics Education Reform Initiative for Teachers)	
Focus on the way mathematics is taught particularly in grades 6-10.	
Schools were selected for SEC based on their participation in the project.	
Science Reform: Project CATS	1995
(Coordinated and Thematic Science)	
Focus on integration of science curriculum particularly grades 6-10.	
Initiative schools were selected for SEC based on the schools with teachers trained in CATS.	
<u>*State Content Standards*</u>	
Instructional Goals and Objectives for Mathematics	1996
Instructional Goals and Objectives for Science	1996

<u>*State Assessments*</u>

NRT: Math, Gr. K-11;
Science: Gr. 3-11

APPENDIX C: Analysis Guide

Sample Sections from Survey

Subject Content: Mathematics

Time on Topic		*Middle School Mathematics Topics*	*Expectations for Students in Mathematics*					
<none>	1	**Number sense / Properties /**	**Memorize**	**Understand Concepts**	**Perform Procedures**	**Analyze / Reason**	**Solve Novel Problems**	**Integrate**
⓪ ① ② ③	10	Place value	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③
⓪ ① ② ③	10	Whole numbers	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③
⓪ ① ② ③	10	Operations	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③
⓪ ① ② ③	10	Fractions	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③
⓪ ① ② ③	10	Decimals	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③
⓪ ① ② ③	10	Percents	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③
⓪ ① ② ③	10	Ratio, proportion	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③
⓪ ① ② ③	10	Patterns	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③
⓪ ① ② ③	10	Real numbers	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③
⓪ ① ② ③	11	Exponents, scientific notation	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③
⓪ ① ② ③	11	Factors, multiples, divisibility	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③	⓪ ① ② ③

INSTRUCTIONAL ACTIVITIES IN MATHEMATICS

Listed below are some questions about what students in the target class do in mathematics. For each activity, pick one of the choices (0, 1, 2, 3) to indicate the percentage of instructional time that students are doing each activity. Please think of an average student in this class, in responding.

What percentage of mathematics instructional time in the target class do students:

NOTE: No more than two '3's , or four '2's should be reported for this set of items.

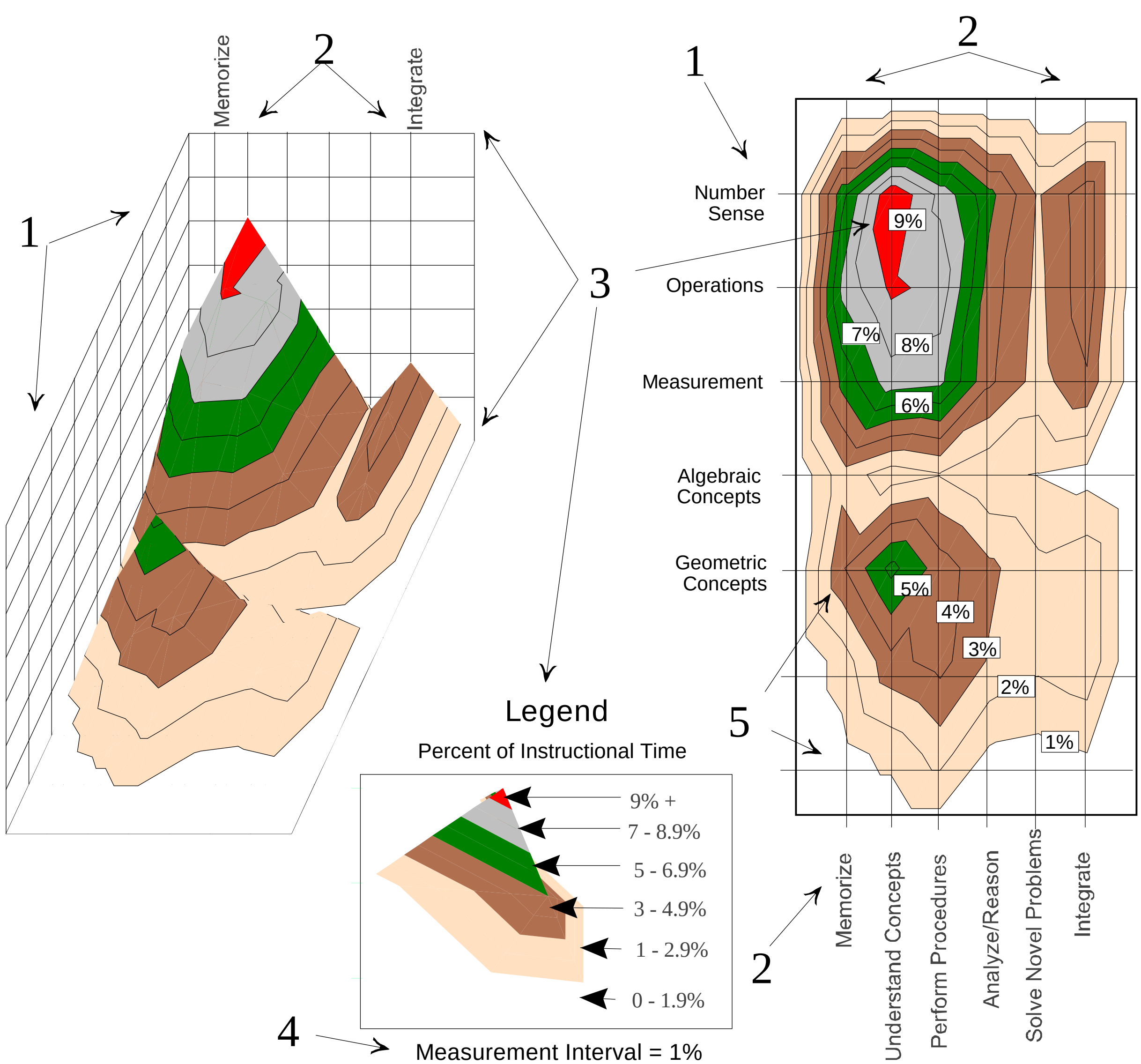
		None	Less than 25%	25% to 33%	More than 33%
34	Watch the teacher demonstrate how to do a procedure or solve a problem.	⓪	①	②	③
35	Read about mathematics in books, magazines, or articles.	⓪	①	②	③
36	Collect or analyze data.	⓪	①	②	③
37	Maintain and reflect on a mathematics portfolio of their own work.	⓪	①	②	③
38	Use hands-on materials or manipulatives (e.g., counting blocks, geometric shapes, algebraic tiles).	⓪	①	②	③
39	Engage in mathematical problem solving (e.g., computation, story-problems, mathematical investigations).	⓪	①	②	③
40	Take notes.	⓪	①	②	③
41	Work in pairs or small groups (non-laboratory).	⓪	①	②	③
42	Do a mathematics activity with the class **outside** the classroom.	⓪	①	②	③
43	Use computers, calculators, or other technology to learn mathematics.	⓪	①	②	③
44	Work individually on assignments.	⓪	①	②	③
45	Take a quiz or test.	⓪	①	②	③

Interpreting Content Maps

Content maps provide a three-dimensional representation of instructional content using a surface area chart which results in a graphic very similar to topographical maps. The grid overlaying each map identifies a list of topics areas (indicated by horizontal grid lines; see 1 below) and six categories of cognitive expectations for students (indicated by vertical lines; see 2 below). The intersection of each topic area and category of cognitive expectation represents a measurement node (see 5 below). Each measurement node indicates a measure of instructional time for a given topic area and category of cognitive expectation based upon teacher reports. The resulting map is based upon the values at each of these measurement nodes. It should be noted that the spaces between each measurement node, that is the surface of the map, are abstractions and are not based upon real data, the image of the map is simply a computer generated graphic based upon the values for each intersecting measurement node. The map display is utilized to portray the third dimension (percent of instructional time; see 3 below) onto this grid utilizing shading and contour lines to indicate the percent of instructional time spent (on average across teachers) for each topic by cognitive expectation intersection.

The increase (or decrease) in instructional time represented by each shaded band is referred to as the measurement interval (see 4 below). To determine the amount of instructional time for a given measurement node, count the number of contour lines between the nearest border and the node, and multiply by the measurement interval.

The graphic at left below displays the three dimensional counterpart of the image represented by the content map displayed on the right. Both graphs indicate that Understanding Concepts related to Number Sense and Operations occupies the majority of time spent on grade four mathematics instruction (9% or more of instructional time over the course of a school year).



Mathematics Scales and Items for each Scale

Math Scales		Reliability Coefficient
Communicating Mathematical Understanding		**0.81**
Q49	Write an explanation to a problem using several sentences.	
Q54	Talk about ways to solve mathematics problems.	
Q57	Work on a writing project where group members help to improve each others' (or the group's) work.	
Q64	Present information to students concerning a mathematical idea or project.	
Q69	Display and analyze data.	
Q81	Individual or group demonstration, presentation.	
Active Learning in Mathematics		**0.91**
Q31	Collect data or information as part of mathematics homework.	
Q36	Collect or analyze data.	
Q38	Use hands-on materials or manipulatives (e.g., counting blocks, geometric shapes, algebraic tiles).	
Q56	Work on an assignment, report, or project that takes longer than one week to complete.	
Q60	Work with hands-on materials such as counting blocks, geometric shapes, or algebraic tiles to understand concepts.	
Q61	Measure objects using tools such as rulers, scales, or protractors.	
Q62	Build models or charts.	
Q63	Collect data by counting, observing, or conducting surveys.	
Q73	Math manipulatives (e.g., pattern blocks, algebraic tiles).	
Q74	Measuring tools (e.g., rulers, protractors, scales).	
Reasoning and Problem Solving		**0.74**
Q19	Field study or out-of-class investigation.	
Q48	Solve novel mathematical problems.	
Q49	Write an explanation to a problem using several sentences.	
Q51	Make estimates, predictions, guesses, or hypotheses.	
Q52	Analyze data to make inferences or draw conclusions.	
Q50	Apply mathematical concepts to real or simulated "real-world" problems.	
Teacher Preparedness for providing an equitable environment		**0.89**
Q100	Teach students with physical disabilities.	
Q102	Teach classes for students with diverse abilities.	
Q103	Teach mathematics to students from a variety of cultural backgrounds.	
Q104	Teach mathematics to students who have limited English proficiency.	
Q105	Teach students who have a learning disability which impacts mathematics learning.	
Q106	Encourage participation of females in mathematics.	
Q107	Encourage participation of minorities in mathematics.	
Q137	Meeting the needs of all students.	

Mathematics Scales and Items for each Scale

Teacher Preparedness for using innovative teaching strategies		**0.92**
Q96	Use/manage cooperative learning groups in mathematics.	
Q97	Integrate mathematics with other subjects.	
Q99	Use a variety of assessment strategies (including objective and open-ended formats).	
Q109	Teach estimation strategies.	
Q110	Teach problem-solving strategies.	
Q112	Teach mathematics with the use of manipulative materials, such as counting blocks, geometric shapes, and so on.	
Professional Collegiality		**0.68**
Q120	I am supported by colleagues to try out new ideas in teaching mathematics.	
Q122	Mathematics teachers in this school regularly share ideas and materials.	
Q123	Mathematics teachers in this school regularly observe each other teaching classes.	
Q126	Most mathematics teachers in this school contribute actively to making decisions about the mathematics curriculum.	
Q127	I have adequate time during the regular school week to work with my peers on mathematics curriculum or instruction.	
Use of Multiple Assessment Strategies		**0.82**
Q79	Extended response item for which student must explain or justify solution.	
Q80	Performance tasks or events (e.g., hands-on activities).	
Q82	Mathematics projects.	
Q83	Portfolios.	
Q84	Systematic observation of students.	
Q101	Help students document and evaluate their own mathematics work.	
Q138	Multiple strategies for student assessment.	
Standards		**0.82**
Q85	Your state's curriculum framework or content standards.	
Q86	Your district's curriculum framework or guidelines.	
Q90	National mathematics education standards.	
Q98	Implement instruction that meets mathematics standards.	
Q133	How to implement state or national content standards.	
Use of Educational Technology		**0.62**
Q43	Use computers, calculators, or other technology to learn mathematics.	
Q67	Use sensors and probes.	
Q68	Collect data or information (e.g., using the Internet).	
Q75	Calculators.	
Q76	Graphing calculators.	
Q139	Educational Technology.	

Science Scales and Items for each Scale

Science Scales		Reliability Coefficient
Communicating Scientific Understanding		**0.72**
Q38	Write about science.	
Q52	Make tables, graphs, or charts.	
Q54	Talk about ways to solve science problems.	
Q56	Write results or conclusions of a laboratory activity.	
Q61	Organize and display the information in tables or graphs.	
Q69	Display and analyze data.	
Active Learning in Science		**0.78**
Q31	Collect data or information about science (as part of science homework).	
Q36	Collect information about science.	
Q39	Do a laboratory activity, investigation, or experiment in class.	
Q41	Work in pairs or small groups.	
Q42	Do a science activity with the class outside the classroom or science laboratory.	
Q47	Use science equipment or measuring tools.	
Q48	Collect data.	
Q49	Change something in an experiment to see what will happen.	
Q57	Work on an assignment, report, or project that takes longer than one week to complete.	
Student Reflection on Scientific Ideas		**0.64**
Q37	Maintain and reflect on a science portfolio of their own work.	
Q58	Work on a writing project or portfolio where group members help to improve each others' (or the group's) work.	
Q60	Ask questions to improve understanding.	
Q63	Discuss different conclusions from the information or data.	
Q64	List positive (pro) and negative (con) reactions to the information.	
Scientific Thinking		**0.8**
Q50	Design ways to solve a problem.	
Q51	Make guesses, predictions, or hypotheses.	
Q53	Draw conclusions from science data.	
Q62	Make a prediction based on the information or data.	
Q65	Reach conclusions or decisions based upon the information or data.	
Teacher Preparedness for providing an equitable environment		**0.81**
Q103	Teach students with physical disabilities.	
Q105	Teach classes for students with diverse abilities.	
Q106	Teach science to students from a variety of cultural backgrounds.	
Q107	Teach science to students who have limited English proficiency.	
Q108	Teach students who have a learning disability which impacts science learning.	
Q109	Encourage participation of females in science.	
Q110	Encourage participation of minorities in science.	
Q136	Meeting the needs of all students.	

Science Scales and Items for each Scale

Teacher Preparedness for using innovative teaching strategies		**0.86**
Q96	Use/manage cooperative learning groups in mathematics.	
Q97	Take into account students' prior conceptions about natural phenomena when planning curriculum and instruction.	
Q100	Integrate science with other subjects.	
Q101	Manage a class of students who are using hands-on or laboratory activities.	
Q102	Use a variety of assessment strategies (including objective and open-ended formats).	
Professional Collegiality		**0.71**
Q117	I am supported by colleagues to try out new ideas in teaching science.	
Q119	Science teachers in this school regularly share ideas and materials.	
Q120	Science teachers in this school regularly observe each other teaching classes.	
Q124	Most science teachers in this school contribute actively to make decisions about the science curriculum.	
Q125	I have adequate time during the regular school week to work with my peers on science curriculum instruction.	
Use of Multiple Assessment Strategies		**0.75**
Q79	Extended response item for which student must explain or justify solution.	
Q80	Performance tasks or events (e.g., hands-on activities).	
Q82	Science projects.	
Q83	Portfolios.	
Q84	Systematic observation of students.	
Q104	Help students document and evaluate their own science work.	
Q137	Multiple strategies for student assessment.	
Standards		**0.67**
Q85	Your state's curriculum framework or content standards.	
Q86	Your district's curriculum framework or guidelines.	
Q90	National science education standards.	
Q98	Implement instruction that meets science standards.	
Q132	How to implement state or national science content standards.	
Use of Educational Technology		**0.6**
Q43	Use computers, calculators, or other technology to learn science.	
Q67	Use sensors and probes (e.g., CBL's).	
Q68	Retrieve or exchange data or information (e.g., using the Internet).	
Q73	Computer/lab interfacing devices.	
Q138	Educational Technology.	

References

American Association for the Advancement of Science (1993). *Benchmarks for science literacy*. New York: Oxford University Press.

Beaton, A.E., Mullis, I.V.S., Martin, M.O., Gonzalez, E.J., Kelly, D.L., & Smith, T.A. (1996). *Mathematics Achievement in the Middle School Years; Science Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study (TIMSS).* Chestnut Hill, MA: Boston College.

Blank, R.K., Langesen, D., Bush, M., Sardina, S., Pechman, E., & Goldstein, D. (1997). *Mathematics and Science Content Standards and Curriculum Frameworks: States Progress on Development and Implementation*. Washington, DC: CCSSO.

Blank, R.K. and Gruebel, D. (1999, biennial report). *State Indicators of Science and Mathematics Education 1999--State Trends and New Indicators from the 1997-98 School Year*. Washington, DC: CCSSO, State Education Assessment Center.

Clune, W.H. (1998). *Toward a Theory of Systemic Reform: The Case of Nine NSF Statewide Systemic Initiatives*. Madison, WI: National Institute for Science Education.

Cohen, David K. (1990). "A Revolution in One Classroom: The Case of Mrs. Oublier." *Educational Evaluation and Policy Analysis; v12n3p[327-45].*

Corcoran, T.B., Shields, P.M., & Zucker, A.A. (1998). *The SSIs and Professional Development for Teachers*. Menlo Park, CA: SRI International.

Council of Chief State School Officers (1998). *How Is Science Taught in Schools? Executive Summary: Five-State Preliminary Study of Use of Enacted Curriculum Surveys in Science*. Washington, DC: Author.

Council of Chief State School Officers (2000). *Summary of Findings from SSI and Recommendations for NSF's Role with States: How NSF Can Encourage State Leadership in Improvement of Science and Mathematics Education*. Washington, DC: Author.

Council of Chief State School Officers (2000). *Using Data on Enacted Curriculum in Mathematics & Science: Sample Results from a Study of Classroom Practices and Subject Content*. Summary Report of the Survey of Enacted Curriculum, with Wisconsin Center for Education Research and 11 States. Washington, DC: Author.

Fuhrman, S.H., ed. (2001). *From Capitol to the Classroom: Standards-based Reform in the States*. Chicago, IL: National Society for the Study of Education and the University of Chicago Press.

Gamoran, A., Porter, A.C., Smithson, J., & White, P.A. (1997, Winter). "Upgrading high school mathematics instruction: Improving learning opportunities for low-achieving, low-income youth." *Educational Evaluation and Policy Analysis, 19*(4), 325-338.

Kahle, J.B., Meece, J.L., & Damnjanovic, A. (1999). "*A* Pocket Panorama of Ohio's Systemic Reform." Summary data from *Bridging the Gap: A Research Study of the Efficacy of Mathematics and Science Education.* Oxford, OH: Miami University.

Klein, S., Hamilton, L., McCaffrey, D., Stecher, B., Rbyn, A., & Burroughs, D. (2000). *Teaching Practices and Student Achievement: Report of First-Year Findings from the "Mosaic" Study of Systemic Initiatives in Mathematics and Science.* Los Angeles: RAND.

Knapp, M.S. (1996). *Between Systemic Reforms and the Mathematics and Science Classroom*. Madison, WI: National Institute for Science Education.

Love, Nancy (2000). *Using Data–Getting Results: Collaborative Inquiry for School-Based Mathematics and Science Reform.* Cambridge, MA: TERC.

Martin, M., Blank, R.K., & Smithson, J. (1996). *Guide for Educators on the Use of Surveys and Data on Enacted Curriculum*. Washington, DC: CCSSO.

Martin, M., Mullis, I., Gonzalez, E., O'Connor, K., Chrostowski, S., Gregory, K., Smith, T., & Garden, R., (2001). *Science Benchmarking Report. TIMSS 1999-Eighth Grade. Achievement for U.S. States and Districts in an International Context*. Chestnut Hill, MA: The International Study Center, Lynn School of Education.

Massell, D., Kirst, M., & Hoppe, M. (No. RB-21-March 1997). *Persistence and Change: Standards-Based Systemic Reform in Nine States*. Philadelphia, PA: CPRE Policy Brief.

Mullis, I., Martin, M., Gonzalez, E., O'Connor, K., Chrostowski, S., Gregory, K., Smith, T., & Garden, R. (2001). *Mathematics Benchmarking Report. TIMSS 1999-Eighth Grade. Achievement for U.S. States and Districts in an International Context*. Chestnut Hill, MA: The International Study Center, Lynn School of Education.

National Center for Education Statistics (1996, 1997, 1998). *Pursuing Excellence: A Study of U.S. Eighth-Grade Mathematics and Science Teaching, Learning, Curriculum, and Achievement in International Context: Initial Findings from the Third International Mathematics and Science Study*. Washington, DC: U.S. Department of Education.

National Center for Education Statistics (1997). *NAEP 1996 Mathematics Cross-State Data Compendium for Grade 4 and Grade 8 Assessment.* Washington, DC: U.S. Department of Education.

National Council of Teachers of Mathematics (1989). *Curriculum and Evaluation Standards for School Mathematics.* Reston, VA: Author.

National Council of Teachers of Mathematics (2000). *Principles and Standards for School Mathematics*. Reston, VA: Author.

National Research Council (1995). *National Science Education Standards*. Washington, DC: National Academy Press.

Porter, A.C. et. al. (1993). *Reform Up Close: An Analysis of High School Mathematics and Science Classrooms: Final Report to the National Science Foundation*. Madison, WI: Wisconsin Center for Education Research.

Porter, A.C. (1998). "The Effects of Upgrading Policies on High School Mathematics and Science." In D. Ravitch (ed.) *Brookings Papers on Education Policy*. Washington, DC: Brookings Institution Press.

Porter, A.C., Garrett, M.S., Desimone, L., Yoon, K.S., & Birman, B. (2000). *Does Professional Development Change Teachers' Instruction: Results from a Three-Year Study of the Effects of Eisenhower and Other Professional Development on Teaching Practice*. Report to the U.S. Department of Education, Office of the Undersecretary, Planning and Evaluation Service, Washington, DC.

Porter, A.C. and Smithson, J. (2001). "Are Content Standards Being Implemented in the Classroom? (2001). "A methodology and Some Tentative Answers." In, *From the Capitol to the Classroom: Standards-Based Reform in the States*. S.H. Fuhrman (ed.) Chicago: National Society for the Study of Education.

Reese, C.M., Miller, K.E., Mazzeo, J., & Dossey, J.A. (1997). *NAEP 1996 Mathematics Report Card for the Nation and the States*. Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Schmidt, W.H., McKnight, C.C., Valverde, G.A., Houang, R.T., & Wiley, D.E. (1996). *Many Visions, Many Aims: A Cross-National Investigation of Curricular Intentions in School Mathematics*. Boston: Kluwer Academic Publishers.

Schmidt, W.H., Raizen, S.A., Britton, E.D., Bianchi, L.J., & Wolfe, R.G. (1996). *Many Visions, Many Aims: A Cross-National Investigation of Curricular Intentions in School Science*. Boston: Kluwer Academic Publishers.

Smithson, J. and Porter, A.C. (1994). *Measuring Classroom Practice: Lessons Learned from Efforts to Describe the Enacted Curriculum*. New Brunswick, NJ: Rutgers University, Consortium for Policy Research in Education.

Smithson, J., Porter, A.C., & Blank, R.K. (1995). *Describing the Enacted Curriculum: Development and Dissemination of Opportunity to Learn Indicators in Science Education*. Washington, DC: CCSSO.

Systemic Research, Inc. (2000). *Urban School: Key Indicators of Science and Mathematics Education*, Volume IV: Overall Progress Report. Norwood, MA: Author.

Webb, N.L. (1999). *Evaluation of Systemic Reform in Mathermatics and Science*. (Synthesis and proceedings of the Fourth Annual NISE Forum, Workshop Report, No. 8). Madison, WI: University of Wisconsin, National Institute of Science Education.

Weiss, I.R. (1994). *A Profile of Science and Mathematics Education in the United States, 1993*. Chapel Hill, NC: Horizon Research, Inc.

Westat and Policy Studies Associates (2000). *The Longitudinal Evaluation of School Change and Performance (LESCP) in Title I Schools, Volume I: Executive Summary and Draft Final Report*. Prepared for the Office of the Under Secretary, U.S. Department of Education. Rockville, MD and Washington, DC: Authors.

Wilson, L.D. and Blank, R.K. (1999). *Improving Mathematics Education Using Results from NAEP and TIMSS*. Washington, DC: CCSSO.

Zucker, A.A., Shields, P.T., Adelman, N.E., Corcoran, T.B., & Goertz, M.E. (1998). *A Report on the Evaluation of the National Science Foundation's Statewide Systemic Initiatives Program*. Menlo Park, CA: SRI International.