DOCUMENT RESUME

ED 477 657 EA 032 563

AUTHOR	Porter, Andrew C.; Smithson, John L.
TITLE	Defining, Developing, and Using Curriculum Indicators. CPRE Research Report Series.
INSTITUTION	Consortium for Policy Research in Education, Philadelphia, PA.
SPONS AGENCY	National Inst. on Educational Governance, Finance, Policymaking, and Management (ED/OERI), Washington, DC.
REPORT NO	CPRE-RR-Ser-048
PUB DATE	2001-12-00
NOTE	44p.
CONTRACT	OERI-R308A60003
AVAILABLE FROM	Consortium for Policy Research in Education, 3440 Market Street, Suite 560, Philadelphia, PA 19104-3325. Tel: 215-573-0700; Fax: 215-573-7914; e-mail: cpre@gse.upenn.edu; Web site: http://www.gse.upenn.edu/cpre/. For full text: http://www.cpre.org/Publications/rr48.pdf.
PUB TYPE	Information Analyses (070) -- Reports - Evaluative (142)
EDRS PRICE	EDRS Price MF01/PC02 Plus Postage.
DESCRIPTORS	Academic Standards; Accountability; Course Content; Curriculum; *Curriculum Based Assessment; *Curriculum Development; *Curriculum Evaluation; Educational Assessment; *Educational Indicators; Elementary Secondary Education; *Instructional Development; Mathematics Curriculum; Science Curriculum; State Standards

ABSTRACT

The purpose of this paper is to report on the progress of work done by the Consortium for Policy Research in Education on developing valid yet efficient measures of instructional content and its relationship to assessment and standards. The goal is to help develop new methodologies to assess the relationships between what is taught and what is desired to be taught. The report focuses on mathematics and science but also touches on language arts and history. It begins with a brief review of the lessons learned in the Reform Up Close study, which discussed the intended versus the enacted curriculum. The report then discusses the central issues involved in defining and measuring curriculum indicators. This is followed by a discussion about using curriculum indicators in school improvement, program evaluation, and informing policy decisions. Considerable attention is paid to new methods for determining alignment among instruction, assessments, and standards. The report concludes with a discussion of the next steps in the development and expansion of curriculum indicators. Appended are comprehensive lists of mathematics and science topics for elementary, middle, and high schools, and details of mathematics cognitive demands. (Contains 4 figures and 26 references.) (WFA)

Defining, Developing, and Using Curriculum Indicators

Andrew C. Porter
John L. Smithson

CPRE Research Report Series
RR-048

December 2001

Consortium for Policy Research in Education
University of Pennsylvania
Graduate School of Education

Contents

List of Figures

Biographies

Andrew Porter is professor of educational psychology and director of the Wisconsin Center for Education Research at the University of Wisconsin-Madison. He has published widely on psychometrics, student assessment, education indicators, and research on teaching. His current work focuses on curriculum policies and their effects on opportunity to learn.

John Smithson is a research associate at the Wisconsin Center for Education Research, where he has worked for the past 10 years on developing indicators of classroom practice and instructional content. He has worked on several federal- and state-funded research projects investigating changes in classroom instruction based upon various reform initiatives.

Acknowledgments

This research was supported by a grant (No. OERI-R308A60003) from the National Institute on Educational Governance, Finance, Policymaking, and Management (Office of Educational Research and Improvement, U.S. Department of Education) to the Consortium for Policy Research in Education (CPRE). The opinions expressed herein are those of the authors and do not reflect the views of the National Institute on Educational Governance, Finance, Policymaking, and Management; the U.S. Department of Education; the Office of Educational Research and Improvement; CPRE; or its institutional members.

Introduction

As education reform efforts have moved toward a standards-based, accountability-driven, and systemically-integrated approach to improving instructional quality and student learning, researchers and policymakers have become increasingly interested in examining the relationship between the curriculum delivered to students and the goals of state and district policy initiatives. Assessing relationships between what is taught and what is desired to be taught has required the development of new methodologies. The purpose of this report is to describe the progress of our work as we have worked to develop valid yet efficient measures of instructional content and its relationships to assessment and standards. We have focused on mathematics and science, but done some work in language arts and history as well. We hope this report is useful to researchers and policymakers who wish to track changes in the content of instruction or to determine relationships between curriculum policies and instructional content.

We begin with a brief review of the lessons learned in the Reform Up Close study, a Consortium for Policy Research in Education (CPRE) project funded by the National Science Foundation, then discuss the central issues involved in defining and measuring curriculum indicators, while noting how our approach has developed over the past 10 years. This is followed by a discussion about using curriculum indicators in school improvement, program evaluation, and informing policy decisions. Considerable attention is paid to new methods for determining alignment among instruction, assessments, and standards. We conclude with a discussion of the next steps in the development and expansion of curriculum indicators.

Defining Measures of the Enacted Curriculum

During the 1990-1992 school years, a team of researchers from the University of Wisconsin, led by Andrew Porter, and Stanford University, led by Michael Kirst, undertook an unprecedented large-scale look behind the classroom door (Porter, Kirst, Osthoff, Smithson, and Schneider, 1993). Incorporating an array of data collection tools, the researchers examined mathematics and science instructional content and pedagogy delivered to students in over 300 high school classrooms in six states. Detailed descriptions of practice were collected, using daily teacher logs, for a full school year in more than 60 of these classrooms.

Interest in descriptions of classroom practice has grown steadily since the early 1990s, particularly as high-stakes tests have become a favored component of state and district accountability programs. In such an environment it is essential that curriculum indicators provide reliable and valid descriptions of classroom practice. Additionally, indicators should be versatile enough to serve the needs of researchers, policymakers, administrators, teachers, and the general public. Our work described here has sought to develop measures and analyses that meet these demands.

Distinguishing the Intended, Enacted, Assessed, and Learned Curricula

Classroom practice is the focal point for curriculum delivery and student learning. So, it is not surprising that policymakers and researchers are interested in understanding the influence of the policy environment (including policies covering standards, assessments, accountability, and professional development) on classroom practice and gains in student achievement. The importance of policies guiding curriculum has led us to expand our conceptual framework to consider the curricular implications.

In the Reform Up Close study, we discussed the *intended* versus the *enacted* curriculum, noting that the intention was that practice (the *enacted* curriculum) should reflect the curriculum policies of the state (the *intended* curriculum). More recently we have come to distinguish the *intended* from the *assessed* curriculum, and the *enacted* from the *learned* curriculum (Porter and Smithson, 2001). These distinctions come from the international comparative studies of student achievement literature that first distinguished among the *intended*, *enacted*, and *learned* curricula (McKnight et al., 1987; Schmidt et al., 1996). One could argue that the *assessed* curriculum is a component of the *intended* curriculum, and the *learned* curriculum an aspect of the *enacted* curriculum. But we have found that these finer distinctions serve an important analytic role in tracing the chain of causality from education legislation to student outcomes.

The Enacted Curriculum

The *enacted* curriculum refers to the actual curricular content that students engage in the classroom. The *intended*, *assessed*, and *learned* curricula are important components of the educational delivery system, but most learning is expected to occur within the *enacted* curriculum. As such, the *enacted* curriculum is arguably the single most important feature of any curriculum indicator system. It has formed the centerpiece of our efforts over the last 10 years; we developed a comprehensive and systematic language for describing instructional content with the *enacted* curriculum in mind.

Descriptions of the *enacted* curriculum still lie at the heart of our work, but we have come to appreciate the importance of looking at the *intended*, *assessed*, and *learned* curricula in combination with the *enacted* curriculum in order to describe the context within which instruction occurs.

The Intended Curriculum

By the *intended* curriculum we refer to such policy tools as curriculum standards, frameworks, or guidelines that outline the curriculum teachers are expected to deliver. These policy tools vary significantly across states, and to some extent, across districts and schools.

There are two important types of information that should be collected when examining the *intended* curriculum. The collected information should include the composition of the curriculum described in policy documents. It is also important to collect measures that characterize the policy documents themselves. For example, how consistent are the policies in terms of curricular expectations? How prescriptive are the policies in indicating the content to be delivered? How much authority do the policies have among teachers? And finally, how much power have the policies in terms of rewards for compliance and sanctions for non-compliance? (Porter, Floden, Freeman, Schmidt, and Schwille, 1988; Schwille et al., 1983). Such policy analyses are distinct from alignment analyses, and both play a critical role in explaining the curriculum delivered to students.

The Assessed Curriculum

Though assessments could be included in the definition of the *intended* curriculum, high-stakes tests play a unique role in standards-based accountability systems, often becoming the criteria for determining success or failure, reward or punishment. Therefore, it is analytically useful to distinguish the *assessed* curriculum (represented by high-stakes tests) from the *intended* curriculum (represented by curriculum standards, frameworks, or guidelines). At a minimum, it can be informative to compare the content in the assessments with the content in the curriculum standards and other policy documents. Such comparisons, in most cases, reveal important differences between the knowledge that is valued and the knowledge that is assessed, differences perhaps due to the limitations of resources and the technologies available for assessing student knowledge. Lack of alignment leads to an almost inevitable tension between the *intended* and the *assessed* curriculum. A curriculum indicator system should be able to reveal this tension and be able to characterize its nature within particular education systems.

The Learned Curriculum

With the advent of standards-based reform and the popularity of accountability systems, student achievement scores are the apparent measure of choice in determining the success of educational endeavors. Just as the *assessed* curriculum is, as a practical matter, restricted to reflecting a subset of the *intended* curriculum, achievement scores represent just a portion of the knowledge that students acquire as a result of their schooling experience. Nonetheless, these measures invariably represent the bottom line for education providers under current reform initiatives.

Achievement scores may provide a reasonable summary measure of student learning, but, alone, they tell us little about the *learned* curriculum. To be useful for monitoring, evaluating, and diagnosing purposes, indicator measures of the *learned* curriculum need to describe the content that has been learned as well as the level of proficiency offered by test scores. In addition, student outcomes should be mapped on the curriculum to provide information about which parts of the curriculum have been learned by large numbers of students and which aspects require increased attention. Several testing services provide skills analyses that tell how well students performed in various content areas. While we applaud such efforts, it is not clear the extent to which such analyses are used by teachers, or the extent to which such analyses employ a sufficiently detailed language to meet the indicator needs of the system.

The Importance of a Systematic and Comprehensive Language for Description

Distinguishing the four components of the curriculum delivery system allows for examination and comparison of the curriculum at different points in the system. Conducting such analyses requires a common language for describing each component of the system. The more systematic and detailed the language, the more precise the comparisons can be (Porter, 1998b).

We have found that the use of a multi-dimensional, taxonomy-based approach to coding and analyzing curricular content can yield substantial analytic power (examples are provided later). The Upgrading Mathematics study conducted by CPRE provides the most compelling evidence to date (Porter, 1998a). Using a systematic and common language for examining the *enacted*, *assessed*, and *learned* curriculum in that study, we were able to demonstrate a strong, positive, and significant correlation (.49) between the content of instruction (that is, the *enacted* curriculum) and student achievement gains (the *learned* curriculum). When we controlled for prior achievement, students' poverty level, and content of instruction (using an HLM approach in our analysis), practically all variation in student learning gains among types of first-year high school mathematics courses was explained (Gamoran, Porter, Smithson, and White, 1997). These results not only attest to the utility of the language, but also the validity of teacher self-reports on surveys to measure the variance in content of instruction.

More recently we have developed procedures for examining content standards and curriculum frameworks (the *intended* curriculum), with an eye toward looking at the level of alignment among the *intended*, *enacted*, and *assessed* curricula (Porter and Smithson, 2001). Such analyses also depend upon the use of a common language across the various curricular components in the system. These analyses provide researchers with alignment measures that are useful in evaluating reform efforts and provide policymakers and administrators with descriptive indicators that are valuable in evaluating reform policies.

There is one more advantage to systematizing the language of description. Thus far, the uses have involved comparing components of the curriculum. Within a given component, one could also use systematic language to gather data from multiple sources in order to validate each source. Here too, the more tightly coupled the language used across collection instruments, the easier the comparison for purposes of validation.

Developing Curriculum Indicators

It is one thing to extol the virtues of valid curriculum indicators, and quite another matter to produce them. Collection instruments vary in their particular measurement strengths and weaknesses. Some instruments, such as observation protocols and daily teacher logs, allow for rich and in-depth language that can cover many dimensions in fine detail. Others, most notably survey instruments, require more concise language that can be easily coded into discrete categories.

In the Reform Up Close study, we employed a detailed and conceptually rich set of descriptors of high school mathematics and science that were organized into three dimensions: *topic coverage*, *cognitive demand*, and *mode of presentation*. Each dimension consisted of a set number of discrete descriptors. *Topic coverage* consisted of 94 distinct categories for mathematics (for example, ratio, volume, expressions, and relations between operations). *Cognitive demand* included nine descriptors: memorize, understand concepts, collect data, order/compare/estimate, perform procedures, solve routine problems, interpret data, solve novel problems, and build/revise proofs. There were seven descriptors for *modes of presentation*: exposition, pictorial models, concrete models, equations/formulas, graphical, laboratory work, and fieldwork. A content topic was defined as the intersection of *topic coverage*, *cognitive demand,* and *mode of presentation*, so the language permitted 94 x 9 x 7 or 5,922 possible combinations for describing content. Each lesson could be described using up to five unique three-dimensional topics, yielding an extremely rich, yet systematic language for describing instructional content.

This language worked well for daily teacher logs and for observation protocols. A teacher or observer, once trained in use of and coding procedures for the language, could typically describe a lesson in about five minutes. Based on this scheme, the data for any given lesson could be entered into the database in less than a minute. Because we employed the same language and coding scheme in our daily logs as in our observation protocols, we were able to compare teacher reports and observation reports for a given lesson.

In developing teacher survey instruments for the study, however, we faced significant limitations. We could not provide a way for teachers to report on instructional content as the intersection of the three dimensions without creating a complicated instrument that would impose undue teacher burden. Instead we employed two dimensions — *content category* and *cognitive demand* — displayed in a matrix format, so that a teacher could report on the relative emphasis placed on each category of *cognitive demand* for each *content category*. Even here we faced limitations. To employ all nine categories of *cognitive demand* would require a matrix of 94 rows and nine columns. To make the instrument easier for teachers to complete, we reduced the *cognitive demand* dimension from nine to four categories. In retrospect, we probably reduced the number of categories of *cognitive demand* too much, but had we used six or seven categories (imposing a greater teacher burden), we still would have faced the problem of translating the levels of detail when comparing survey results to log results. As a result, we could make very precise comparisons between observations and teacher reports, but we had less precision in comparing teacher logs and teacher surveys. Since the Reform Up Close study, we have reached a compromise of six categories of *cognitive demand.* Although we have not used teacher logs since the Reform Up Close study, we have employed observation protocols using these same six categories.

Content vs. Pedagogy

Using survey instruments, we were able to collect information on *modes of presentation* and other pedagogical aspects of instruction, but did not integrate the information with *topic coverage* and *cognitive demand* in a way to report on the intersection of the three dimensions. If one believes as we do that the interaction of content and pedagogy most influences achievement, then this is a serious loss to the language of description. Of course, there is much more to pedagogy than the *mode of presentation*. Indeed, the concepts of content and pedagogy tend to blur into one another. For that reason, we would ideally define instructional content in terms of at least three dimensions (see discussion below). But, in developing the survey instruments for the Reform Up Close Study, our reporting format required a two-dimensional matrix, thus we had to choose between *cognitive demand* and *mode of presentation*.

We have not lost interest in pedagogy and other aspects of the classroom that influence student learning. For our work with the State Collaborative on Assessment and Student Standards, we developed two distinct sets of survey instruments — one focused on instructional content and the other focused on pedagogy and classroom activities. In a sense, this de-coupled pedagogy from the taxonomic structure we use to describe content, however, and descriptions of content have best explained student achievement.

While we have focused our attention of late on a two-dimensional construct of content, we are still considering the introduction of a third, more pedagogically-based dimension into the language. One possibility is using multiple collection forms crossed on rotated dimensions to allow selection of interactions of interest for a particular data collection effort, while still maintaining a systematic and translatable connection to the larger multi-dimensional model of description. In this way, one might investigate *modes of presentation* by categories of *cognitive demand*, or alternately, *topics covered* by *mode of presentation*, depending upon the descriptive needs of the investigation.

For example, in the language arts and history survey instruments we developed for CPRE's Measurement of the Enacted Curriculum project, we provided a rotated matrix that asked teachers to report on the interaction between category of *cognitive demand* and *mode of presentation* (see Figure 1). In a small, initial pilot involving three elementary language arts teachers and three middle school history teachers, the teachers reported no difficulty in using the rotated matrix design. The results showed fairly dramatic differences between teacher reports, even when teaching the same subject at the same grade level in the same school. We have not yet employed this strategy on a large scale (or with the mathematics or science versions of our instruments), but it may prove to be a useful strategy for investigating particular questions.

Issues in Developing a Curriculum Indicator System

There are several problems in defining indicators of the content of instruction that must be solved (Porter, 1998b).

Do We Have the Right Language?

Getting the right grain size. One of the most challenging issues in describing the content of instruction is deciding what level of detail of description is most useful. Too much or too little detail both present problems. For example, if description were at the level of only distinguishing math from science, social studies, or language arts, then

Figure 1. Example of Rotated Matrix

SECTION III
Instructional Activities

In this section you are asked to provide information on the relative amount of instructional time devoted to various ways in which instruction is presented to the target class during Language Arts instruction. As with the content section just completed, there are two steps involved in responding to this section:

1. In the table that follows, you are asked to first determine the percent of instructional time spent on each mode of presentation listed. Refer to the "Relative Time Codes" below for indicating the percent of instructional time spent using each mode. Assume that the entire table totals 100%. An "other" category is provided in case there is an important mode of presenting instructional material that is not included in the table. If you indicate a response for the "other" category, please identify the additional means of instructional presentation in the space provided.

2. After indicating the percentage of time spent on each mode of presentation with the target class, use the columns to the right of each mode of presentations to indicate the relative emphasis on each of the seven performance goals identified. Refer to the "Performance Goal Codes" below for indicating your response.

Relative Time Codes: 0 = None 1 = less than 10% 2 = 10% to 25% 3 = 25% to 49% 4 = more than 50%
Performance Goal Codes: 0 = Not a performance goal for this topic; 1 = less than 25%; 2 = 25% to 33%; 3 = more than 33%

Your Performance Goals for Students

Relative Time on Task	15	Modes of Presentation	Memorize, Recall	Understand Concepts	Communicate, Empathize	Investigate	Analyze	Evaluate	Integrate
⓪ ① ② ③ ④	1501	Whole class lecture	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
⓪ ① ② ③ ④	1502	Teacher demonstration	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
⓪ ① ② ③ ④	1503	Individual student work	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
⓪ ① ② ③ ④	1504	Small group work	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
⓪ ① ② ③ ④	1505	Test, quizzes	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
⓪ ① ② ③ ④	1506	Field study, out-of-class investigations	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
⓪ ① ② ③ ④	1507	Whole class discussions	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
⓪ ① ② ③ ④	1508	Student demonstrations, presentations	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
⓪ ① ② ③ ④	1509	Homework done in class	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
⓪ ① ② ③ ④	1510	Multi-media presentations (e.g. film, video, computer, internet)	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
⓪ ① ② ③ ④	1511	Whole class simulations (e.g. role-play, games, real-world simulations	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
⓪ ① ② ③ ④	1512	Other: ______________	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

all math courses would look alike. Nothing would be learned beyond what was already revealed in the course title. On the other hand, if content descriptions identify the particular exercises on which students are working, then all mathematics instruction would be unique. At that level of detail, trivial differences would distinguish between two courses covering the same content.

One issue related to grain size is how to describe instruction that does not come in neat, discrete, mutually exclusive pieces. One particular instructional activity may cover several categories of content and involve a number of cognitive abilities. The language for describing the content of instruction must be capable of capturing the integrated nature of scientific and mathematical thinking.

Getting the right labels. The labels used in describing the content of instruction to denote the various distinctions are extremely important. Ideally, labels are chosen that have immediate face validity for all respondents so that questionnaire construction requires relatively little elaboration beyond the labels themselves. Instrumentation where the language has the same meaning across a broad array of respondents is needed for valid survey data.

Some have suggested that our language would be improved if the terms and distinctions better reflected the reform rhetoric of the mathematics standards developed by the National Council of Teachers of Mathematics (2000) or the science standards of the National Research Council (1996). But the purposes of the indicators described here are to characterize practice as it exists, and to compare that practice to various standards. For these purposes, a reform-neutral language is appropriate. Still, one could argue that the language described here is not reform-neutral but pro status quo. Ideally, the language should be translatable into reform language distinctions so comparison to state and other standards is possible.

Another way to determine the adequacy of the content language is to ask teachers for feedback. As we have piloted our instruments with teachers, their feedback has been surprisingly positive. In general, teachers have found the language to be sufficiently detailed to allow them to describe their practice, although they have suggested (and in some cases we have adopted) changing the terminology for a particular topic or shifting a topic to a different grade level. Some teachers have commented that their instruction is more integrated than the discrete categories of content and cognitive demand that we employed, but the teachers typically were able to identify the various components of their instructional content with the language we have developed.

Getting the right topics. Have we broken up the content into the right sets of topics? Since the Reform Up Close study, we have revised the content taxonomy several times. In each revision the *topic coverage* categories dimension was re-examined, and in some cases, re-organized, yet the resulting topics and organizing categories remain quite similar to the Reform Up Close study framework. We believe we have established a comprehensive list of topics, particularly for mathematics and science (see Appendix A and B), but there are other approaches to organizing topics that may prove useful as well.

One alternative framework is in the beginning stages of development, under the auspices of the Organisation for Economic Co-operation and Development as part of their plan for a new international comparative study of student achievement. Big ideas — such as chance, change and growth, dependency and relationships, and shape — are distinguished in this framework. This is a very interesting way of dividing mathematical content and very different from our approach discussed here. Still, if the goal is to create a language for describing practice, practice is currently organized along the lines of algebra, geometry, and measurement, not in terms of big ideas. Perhaps practice should be reformed to better reflect these big ideas, but that has not yet happened.

Getting the right cognitive demands. When describing the content of instruction, it is necessary to describe both the particular content categories (for example, linear algebra or cell biology) and the cognitive activities that engage students in these topics (such as memorizing facts or solving real-world problems). A great deal of discussion has centered on how many distinctions of cognitive demand there should be, what the distinctions should be, and how they should be defined. The earliest work focusing on elementary school mathematics had just three distinctions: conceptual understanding, skills, and applications (Porter et al., 1988). The Reform Up Close study of high school mathematics and science (Porter et al., 1993) had nine distinctions used for both mathematics and science: memorize facts/definitions/equations; understand concepts; collect data (for example, observe or measure); order, compare, estimate, approximate; perform procedures, execute algorithms, routine procedures (including factoring, classifying); solve routine problems, replicate experiments or proofs; interpret data, recognize patterns; recognize, formulate, and solve novel problems or design experiments; and build and revise theory, or develop proofs.

Since then, the cognitive demand categories have undergone several revisions, mostly minor, and generally settling on six categories. The most recent revisions, while similar to previous iterations, are more behaviorally defined, indicating the knowledge and skills required of students, and providing examples of the types of student behaviors that reflect the given category. We believe that these more detailed descriptions of the *cognitive demand* categories will assist teachers in describing the cognitive expectations they hold for students within particular content categories (see Appendix C).

One language or several? Another related issue concerns the need for different languages to describe the *topic coverage* of instruction at different grade levels within a subject area, or to describe different subjects within a given grade level. Similarly, the categories of *cognitive demand* may need to vary by subject and grade level. Of course, the more the language varies from grade to grade, or subject to subject, the more difficult it is to make comparisons, or aggregate across subjects and grade levels. For that reason we have tried, where practical, to maintain a similar set of categories across grade levels, and to a lesser extent, across subjects. In the Reform Up Close study (Porter et al., 1993), we used the same categorical distinctions to describe *cognitive demands* for both mathematics and science. Obviously, the *topic coverage* categories differed between the two subjects, but we hoped that using the same *cognitive demand* categories would allow some comparisons between mathematics and science.

More recently, the categories of *cognitive demand* have diverged for mathematics and science (See Figure 2). In developing a prototype language for language arts and history, subject specialists have suggested a quite different set of categories for *topic coverage* categories and *cognitive demand*. Thus, the tendency appears to be moving from a single language to multiple languages to describe instructional content. Given the differences across subjects, this may be inevitable, but it does make aggregation of data and comparisons across subjects more difficult.

The Possibility of a Third Dimension

Throughout the development of questionnaires for surveying teachers on the content of their instruction, we have considered adding a third dimension to the content matrix. In the Reform Up Close study, we referred to this third dimension as *mode of presentation*. The distinctions included: exposition — verbal and written, pictorial models, concrete models (for example, manipulatives), equations or formulas (for example, symbolic), graphical, laboratory work, and fieldwork. We have tried different categories of *modes of presentation* at different times. However, *mode of presentation* proved difficult to integrate into the survey version of the taxonomy (as discussed above) and when employed, it did not appear to add power to the descriptions provided by topics and cognitive demand. *Mode of presentation* has not correlated well with other variables, or

Figure 2. Changes in Categories of Cognitive Demand Over Time

Reform Up Close (1989)	Upgrading Mathematics (1993)	Surveys of the Enacted Curriculum (1999)	Data on the Enacted Curriculum (2001)
Mathematics & Science Memorize facts/definitions/ equations Understand concepts Collect data (e.g., observe, measure) Order, compare, estimate, approximate Performing procedures: execute algorithms/routine procedures, classify Solve routine problems, replicate experiments, replicate proofs Interpret data, recognize patterns Recognize, formulate, and solve novel problems/ design experiments Build & revise theory/ develop proofs	***Mathematics & Science*** Memorize facts Understand concepts Perform procedures/ Solve equations Collect/interpret data Solve word problems Solve novel problems	***Mathematics*** Memorize *(w/ 3 descriptors)* Understand concepts *(w/ 5 descriptors)* Perform procedures *(w/ 7 descriptors)* Analyze/reason *(w/ 6 descriptors)* Solve novel problems *(w/ 2 descriptors)* Integrate *(w/ 3 descriptors)* ***Science*** Memorize *(w/ 3 descriptors)* Understand concepts *(w/ 4 descriptors)* Perform procedures *(w/ 5 descriptors)* Conduct experiments *(w/ 5 descriptors)* Analyze information *(w/ 3 descriptors)* Apply concepts & make connections *(w/ 4 descriptors)*	***Mathematics*** Memorize facts, definitions, formulas *(w/ 3 descriptors)* Communicate understanding of mathematical concepts *(w/ 5 descriptors)* Perform procedures *(w/ 7 descriptors)* Conjecture, generalize, prove *(w/ 7 descriptors)* Solve non-routine problems/Make connections *(w/ 4 descriptors)* ***Science*** Memorize facts, definitions, formulas *(w/ 3 descriptors)* Communicate understanding of science concepts *(w/ 4 descriptors)* Perform procedures Conduct investigations *(w/ 8 descriptors)* Analyze information *(w/ 4 descriptors)* Apply concepts & make connections *(w/ 4 descriptors)*

with student achievement gains. Perhaps the problem is its definition, or perhaps *mode of presentation* is not really useful.

A related dimension that has been suggested is *mode of representation*. This dimension would differentiate the manner in which subject matter is *represented* as part of instruction (for example, written, symbolic, or graphic representation). We have not tried to employ this additional dimension thus far, primarily due to considerations of teacher burden.

Teacher pedagogical content knowledge is another dimension that we have not investigated ourselves, but observed with interest the work of others. Our interest in *pedagogical content knowledge* concerns the effect it may have on teachers' descriptions of their instruction. Looking at the reports provided by teachers over the past 10 years, we see a trend toward a more balanced curriculum. Teachers in the early 1990s were reporting a great deal of focus on procedural knowledge and computation, with very little novel problem-solving or real-world applications. Today, teachers report more activities focused on more challenging cognitive demands, although procedural knowledge and computation continue to dominate in mathematics. But we do not know how well reports from teachers with less experience and knowledge will compare to the reports of teachers with a greater depth of content knowledge. One might expect that teachers with more content knowledge would report less time spent on the more challenging cognitive domains because they understand the difficulties in engaging students in cognitively challenging instruction. Novice teachers, by comparison, might over-report the time spent on challenging content because of their under-appreciation of what is entailed in providing quality instruction and ensuring student engagement in non-routine problem-solving, applying concepts, and making connections. The addition of a dimension that measures teacher content knowledge might provide a means of explaining variation across teacher responses that could strengthen the predictive power of curriculum indicator measures on student achievement gains.

Who Describes the Content?

From the perspective of policy research, teachers are probably the most important respondents, because teachers make the ultimate decisions about what content gets taught to which students, when it is taught, and according to what standards of achievement. Curriculum policies, if they are to have the intended effect, must influence teachers' content decisions. Since the period of instruction to be described is long (at least a semester), teachers and students are the only ones likely to be in the classroom for the full period. Because content changes from week to week, if not day to day, a sampling approach by observation or video simply will not work. Video and observation have been used to good effect in studying pedagogical practice, but have worked well only when those practices have been so typical that they occur in virtually every instruction period. However, some pedagogical practices are not sufficiently stable to be well studied, even with a robust sampling approach (Shavelson and Stern, 1981).

Students could be used as informants reporting on the content of their instruction. One advantage of using students is that they are less likely than teachers to report intentions rather than actual instruction. A danger of using students as respondents is that their ability to report on the content of instruction may be confounded by their understanding of that instruction. The reporting of struggling students on instructional content might be incomplete or inaccurate due to their misunderstanding or lack of recall. We conclude that it is more useful to look to teachers for an accounting of what was taught, and to students for an accounting of what was learned.

Response Metric

For respondents to describe the content of instruction, they must be presented with accurate distinctions in type of content, as discussed above. They also need an appropriate metric for reporting the amount of emphasis placed on each content alternative. The ideal metric for emphasis is time: How many instructional minutes were allocated to a particular type of content? This is a metric that facilitates comparisons across classrooms, types of courses, and types of student populations. But reporting number of instructional minutes allocated to a particular type of content over an instructional year is no easy task. The challenge lies in getting a response metric as close as possible to the ideal, in a manner which respondents find manageable and can use with accuracy. Common response metrics include: number of hours per week (in a typical week), number of class periods, frequency of coverage or focus (for example, every day or every week), and relative emphasis. The advantages of these metrics are that they are relatively easy to respond to (particularly for large time frames such as a semester or year) and they are fairly concrete time frames (class period, day, or week). Their major disadvantage is that they yield a fairly crude measure of instructional time.

We settled on a middle approach, using a combination of number of class periods and relative time emphasis in order to calculate the percent of instructional time for a given time period. The *topic coverage* component of the content language is based on number of class periods. The response metric is: (0) not covered, (1) less than one class or lesson, (2) one to five classes or lessons, and (3) more than five classes or lessons. For each *topic covered*, respondents report the relative amount of time spent emphasizing instruction focused on each category of *cognitive demand*. These response metrics are: (0) not a performance goal for this topic, (1) less than 25 percent of time on this topic, (2) 25 to 33 percent of time on this topic, and (3) more than 33 percent of time on this topic. This may at first appear to be a rather skewed and perhaps peculiar metric, but we have found that it divides the relative time spent on a topic into chunks of time that teachers can easily use. Using these response metrics, we are able to calculate an overall percentage of instructional time for each cell in the two-dimensional content matrix (*topic coverage* by *cognitive demand*). We can convert the information on the frequency and length of class periods, if desired, into relative measures of instructional minutes.

How Frequently Should Data Be Collected?

A tension exists between requiring more frequent descriptions to obtain reporting accuracy, which is expensive, and less frequent descriptions covering longer periods of instruction (say, a semester or full school year) which is less expensive and less burdensome, but may be less accurate as well. What frequency of reporting has an acceptable cost but still provides acceptable accuracy? We have used daily logs, weekly surveys, twice-yearly surveys, and a single year-end survey. In the Reform Up Close study, we found good agreement when matching daily logs (aggregated over a school year) to a single year-end survey (average correlation of .54 for mathematics *topic coverage* and .67 for science *topic coverage*). Correlations for the *cognitive demand* categories were difficult to calculate because of differences in log and survey response categories: there were 10 *cognitive demand* categories for the daily logs, but only four categories for the surveys. For the two *cognitive demand* categories (memorize and solve novel problems) that were defined the same for teacher logs and surveys, the correlations were .48 and .34 respectively. Other comparisons between log data and survey data revealed similar results: the average correlation for *modes of instruction* was .43 and the average correlation on *reports of student activities* was .46 (Smithson and Porter, 1994). While these measures are not ideal (and further work comparing log and survey data is needed), they indicate that descriptions of instruction based on a one-time, year-long report do provide descriptions of instruction that resemble descriptions gathered on a daily basis over a full school year. If money, human resources, and teacher burden are no object, daily reports of practice will yield more accurate descriptions of practice. As a more practical matter, however, large-scale use of daily logs is not a viable option. More work is needed to determine the best time frame for gathering teacher reports, but we believe that a single year-long survey instrument is adequate for many of the descriptive and analytic needs for program evaluation. In the CPRE Upgrading Mathematics study, for example, we found that end-of-semester surveys for content descriptions correlated .5 with student achievement gains.

Determining the instructional unit of time that should be described could also affect decisions about the frequency of reporting. At the high school level, the unit might be a course, but some courses last for two semesters while others for only a single semester. Alternatively, the unit might be a sequence of courses used to determine, for example, what science a student studies in a three-year sequence of science courses. At the elementary school level, policymakers are typically interested in the school year or a student's entire elementary school experience (or at least the instruction experienced up to the state's first assessment). Using the semester as the unit of measure seems a reasonable compromise between daily and year-long reporting, but until more work is done to establish the relative utility of semester and year-long reports, we prefer year-long reports, due to cost concerns.

Validating Survey Data

In most efforts to describe the *enacted* curriculum, teachers have reported on their own instruction. The use of teacher self-reported data, however, raises important questions about teacher candor and recall, as well as the adequacy of the instrumentation to provide useful descriptions and teacher familiarity and fluency in the language.

Teacher candor is likely the most frequently raised concern with respect to self-reported data, but probably the least problematic, as long as teacher responses are not used for teacher evaluations. When not linked to rewards or sanctions, teacher descriptions of practice have generally been consistent with the descriptions of practice provided by other sources, whether those sources are findings from other research, classroom observations, or analyses of instructional artifacts (Smithson and Porter, 1994; Burstein et al., 1995; Porter, 1998a; Mayer, 1999).

Even a teacher's best efforts to provide accurate descriptions of practice, however, are constrained by the teacher's ability to recall instructional practice and the extent to which teachers share a common understanding of the terms used in the language of description. Therefore, it is important to conduct analyses into the validity of survey measures in order to increase confidence in survey data. We and others have undertaken several approaches to examine the validity of survey reports.

For the Reform Up Close study, independent classroom observations were conducted on selected days of instruction. When we compared observers' descriptions and the teachers' self reports, we found strong agreement between the teachers and observers (.68 for fine-grain *topic coverage* and .59 for categories of *cognitive demand*), and fair agreement between teacher logs and teacher questionnaires, as discussed above (Smithson and Porter, 1994). Burstein and McDonnell used examples of student work (such as assignments, tests, and projects) to serve as benchmarks and to validate survey data. They found good agreement between these instructional artifacts and reports of instruction (Burstein et al., 1995), but noted the importance of carefully defined response options for survey items, as we have (Smithson and Porter, 1994). Researchers at the National Center for Research on Evaluation, Standards, and Student Testing are also developing indicator measures based on student work (Aschbacher, 1999). Others have used a combination of interviews and classroom observations to confirm our findings on validating survey reports (Mayer, 1999). All of these attempts to validate survey reports have yielded promising results. Still, it is important to continue validating survey measures through the use of alternative data sources, in particular to establish good cost/benefit comparisons for various reporting periods and collection strategies.

Conducting Alignment Analyses

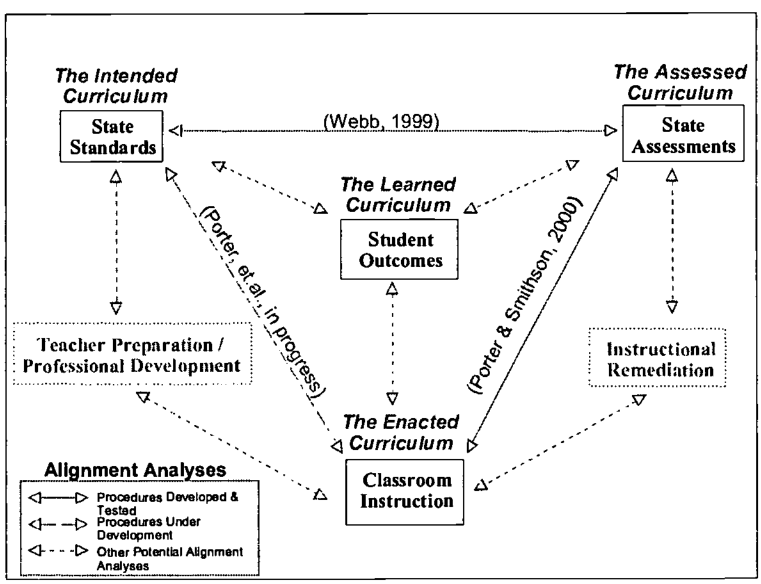
To date, two distinct methodologies for conducting alignment analyses have been developed and field-tested (Porter and Smithson, 2001; Webb, 1999). While there are important differences between the two procedures, they share a basic structure that provides a general picture of how to conduct alignment analyses of standards-related policies and practices.

Both approaches are based on collection of comparable descriptions for two selected components of the standards-based system (see Figure 3). Because these descriptions are the basis of the analysis that results in quantitative measures, the language used in describing those components is a critical element in the process. The language should be systematic, objective, comprehensive, and informative on three dimensions: categorical congruence, breadth, and depth (Webb, 1997).

Alignment Criteria

The most straightforward criteria to use in measuring alignment would be something along the lines of what Webb (1997) calls "categorical concurrence." Here an

Figure 3. Developed and Potential Alignment Analyses



operational question is, for example, "Does this assessment item fit one of the categories identifiable in the standards being employed?" If the answer is yes, we say the item is aligned. If we answer yes for every such item in a state assessment, using categorical concurrence, we say that the assessment is perfectly aligned to the standards.

One does not have to give this approach much consideration before seeing some significant shortcomings in its use as a measure of alignment. For one thing, an assessment that focused exclusively on one standard to the exclusion of all the rest would be equally well aligned as an assessment that equally represented each standard. An alignment measure based on categorical congruence alone could not distinguish between the two, although the two tests would be dramatically different in the range of content assessed.

This leads to a second criteria that would improve the theoretical construct of alignment: a range or breadth of coverage. An assessment can test only a portion of the subject matter that is presented to students. It is important then that assessments used for accountability purposes represent a balance across the range of topics in which students are expected to be proficient. An alignment measure that speaks to range of coverage allows investigation into the relationship between the subject matter range identified in the content standards and the range of topics represented by a particular test.

Breadth of coverage is an improvement over simple categorical congruence, but it is becoming increasingly clear that depth of coverage represents an important ingredient for student success on a given assessment (Gamoran, Porter, Smithson, and White, 1997; Porter, 1998b). Depth of coverage refers to the performance goals or cognitive expectations of instruction, and provides a third dimension to include in calculating an alignment measure.

Alignment Procedures

Two approaches for measuring alignment use some version of these three criteria in their implementation. The two procedures vary in key ways, but both use a two-dimensional grid to map content descriptions for system components in a common, comparable language. Comparisons are made between the relevant cells on the two maps in order to measure the level of agreement between the system components. The results of these quantitative comparisons produce the alignment indicators that can inform policymaking and curricular decision-making.

The first approach simply takes the absolute value of the difference between percent of emphasis on a topic, say, in a teacher's instruction and on a test. The index of alignment is equal to $1-((\Sigma|y-x|)/2)$ where Y is the percent of time spent in instruction and X is the percent of emphasis on the test. The sum is all topics in the two-dimensional grid. The index is 1.0 for perfect alignment and zero for no alignment. This index is systematic in content in that both situations — content not covered on the test but covered in instruction *and* content not covered in instruction but covered on the test — lead to lack of alignment.

The second approach to measuring alignment is a function of the amount of instructional emphasis on topics that are tested. There are two pieces to this second index: one is the percent of instructional time spent on tested content; and the other, for topics that are tested and taught, the match in degree of emphasis in instruction and on the test.

The first index is best suited to looking at consistency among curriculum policy instruments and the degree to which content messages of the policy instruments are reflected in instruction. The second index is the stronger predictor of gains in student achievement.

Using Curriculum Indicators

There are many possible uses of curriculum indicators (Porter, 1991). One use is purely descriptive: what is the nature of the educational opportunity that schools provide? A second use is as an evaluation instrument for school reform. A third use is to suggest hypotheses about why school achievement levels are not adequate.

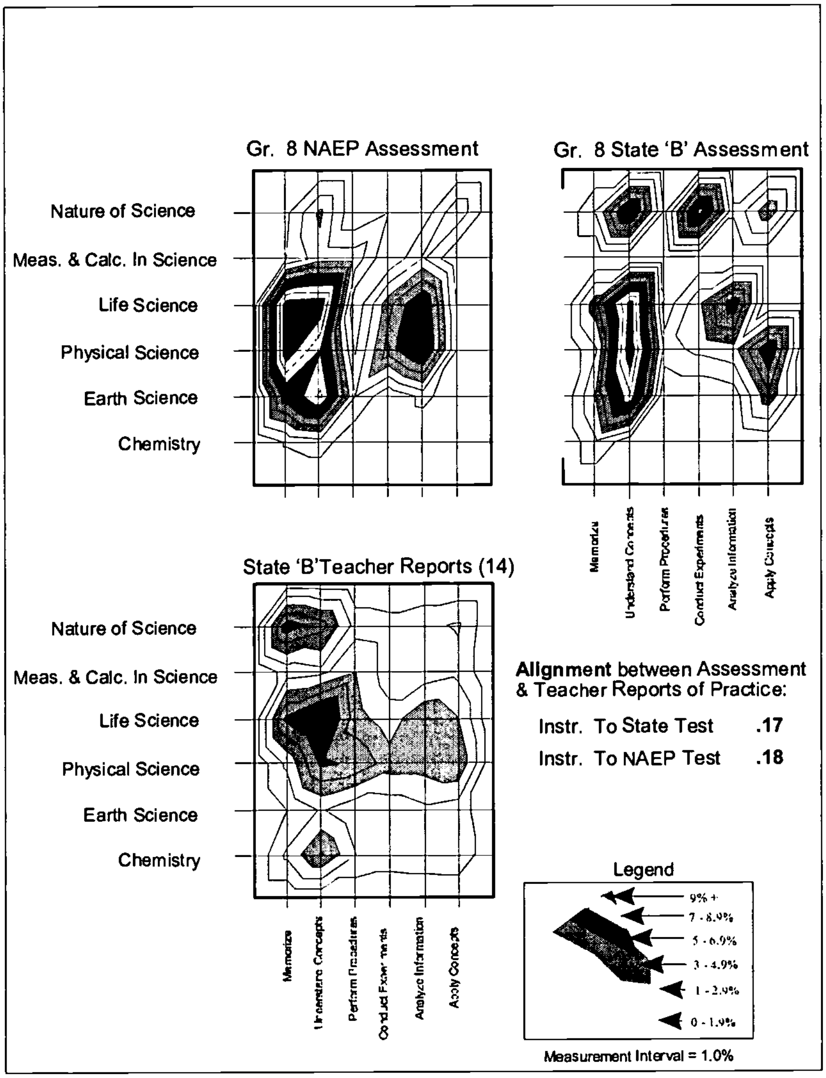
State, District, and School Use

States, particularly those with high-stakes tests or strong accountability policies, have a vested interest in curriculum indicators. Such indicators are crucial in determining the health of the system and measuring the effects of policy initiatives on instruction. In addition, many states must be prepared to demonstrate to a court that students are provided the opportunity to learn the material on which they are assessed (Porter at al, 1993; Porter, 1995).

An indicator system that can provide a picture of the instructional content and classroom practices enacted in a state's schools provides an important descriptive means for monitoring practice. In addition to monitoring their reform efforts, states are interested in providing districts and schools with relevant information to better inform local planning and decision-making. Districts often have curriculum specialists or resource people who value indicator measures for their schools, not only to assist in planning professional development opportunities, but also in some cases to serve as the basis for the professional development activities. Curriculum indicator data at the classroom level can facilitate individual teacher reflection, either during data collection (as reported by teachers in piloting the instruments) or in data reporting (as we have seen in our current work with four urban school districts).

Of particular interest to district and school staff are content maps that juxtapose images of instructional content and a relevant state or national assessment (see Figure 4). The two space of the map represents *topic coverage* categories by *cognitive demand.* Degree of emphasis on topics in the two space is indicated by darkness of color (for example, white indicates content receiving no emphasis). Such graphic displays assist teachers in understanding the scope of particular assessments as well as the extent to which particular content areas may be over- or under-emphasized in their curriculum. We are currently developing procedures to provide similar displays of the *learned* curriculum that teachers could use in determining the content areas where their students need most help.

Figure 4. Grade Eight Science Alignment Analysis



Policy Analysis

The value of curriculum indicators in policy analysis is three-fold. First, indicators of the curriculum provide a mechanism for measuring key components of the standards-based system. This allows careful examination of the relationship between system components in order to determine the consistency and prescriptiveness of policy tools. Secondly, descriptions of curricular practice provide a baseline and means for monitoring progress or change in classroom practice. The effects of policy strategies on instruction can be examined and their efficacy assessed. Finally, if there is interest in attributing student achievement gains to policy initiatives, curriculum alignment indicators provide information on the important intervening variable of classroom instruction.

Analyses of *horizontal* alignment, for example, allow an investigator to examine the degree of consistency among policy tools employed *within* a level of the system (such as the state level). Analyses of *vertical* alignment by contrast describe consistency *across* levels of the system for a given type of policy instrument (say, content standards).

In addition, alignment measures provide a means for holding instructional content constant when examining the effects of competing pedagogical approaches. While many in the educational community are looking for evidence to support the effectiveness of one or another pedagogical approach in improving test scores, obtaining such evidence has proven difficult, we would argue, in large part because the content of instruction has not been controlled. This approach would reconceptualize earlier process-product research on teaching, changing from a search for pedagogical practices that predict student achievement gains to a search for pedagogical practices that predict student achievement gains after first holding constant the alignment of the content of instruction with the content of the achievement measure. Alignment analyses provide such a control, and thus have the potential to permit examination into the effects of competing pedagogical approaches to instruction.

Alignment analyses can also serve to validate teacher reports of practice. If alignment indices based upon teacher reports and content analyses of assessments succeed in predicting student achievement gains as they did in the Upgrading Mathematics Study (Gamoran et al., 1997; Porter, 1998b), then the predictive validity of those teacher reports has been established.

Next Steps for Curriculum Indicators

Interest in curriculum indicators has grown steadily during the past decade as standards-based, systemic reform initiatives have placed greater attention on what occurs behind the classroom door. Significant steps have been taken in the development of instruments and analyses to support an indicator system for describing and comparing the *enacted*, the *intended*, the *assessed,* and the *learned* curricula. Still, some of the most exciting work with curriculum indicators lies just on the horizon of future developments and next steps.

Language and Instrumentation

While a good deal of progress has been made in developing and refining instruments for mathematics and science, we see a variety of opportunities for further development that could increase the quality and scope of the instruments available for curriculum and policy analyses.

Expansion of Subject Areas

To date, the greatest amount of work on curriculum indicators has focused on mathematics and science (Council of Chief State School Officers, 2000; Blank, Kim, and Smithson, 2000; Kim, Crasco, Smithson, and Blank, 2001; Mayer, 1999; Porter, 1998b; Schmidt, McKnight, Cogan, Jakwerth, and Houang, 1999). Draft instruments for language arts and history have been developed as part of the CPRE-funded Measurement of the Enacted Curriculum project, but further field testing is needed before these instruments are ready for use. Additionally, CPRE researchers at the University of Michigan are working on instrumentation for mathematics and reading.

The extent to which instrumentation for other subject areas will be developed will likely follow the emphases states place upon subject areas, especially in their assessment programs. At the moment, mathematics, language arts, and science receive the greatest amount of attention; it is precisely these instruments which have undergone or are undergoing the most development.

Expanding the Taxonomy

As discussed previously, there are other dimensions of the curriculum and instructional practice that are worthy of investigation. Whether a category such as *modes of presentation* or *modes of representation* or *teacher pedagogical content knowledge* would best serve descriptive and analytic needs is unclear and deserves investigation.

The primary advantage of building additional dimensions into the taxonomy is that it allows for a broader descriptive language that could facilitate both collaborative work and meta-analyses for studies with intersecting areas of interest. Further, such additions may increase the analytic power of the resulting measures. While measurement of more than two dimensions is difficult in semester and year-long survey reports, the use of rotated matrices or electronic instrumentation (see discussion below) may provide mechanisms for collecting integrated measures on multiple dimensions. Moreover, instruments such as observation protocols and teacher logs are even more flexible in measuring multiple dimensions, and may serve important descriptive, analytic, and professional development needs where reports based on time frames shorter than a semester are of interest.

Developing Electronic Instrumentation

Data collection and entry are seldom easy, and typically take up the bulk of the logistical activities of research staff. Electronic submissions of data offer an opportunity to dramatically reduce the need for human and paper resources. Electronic data submissions are likely to face many of the same challenges as paper with respect to response and completion rates, but the streamlining of data collection and entry, and the potential for quick and substantive feedback to users, offers an opportunity too valuable to ignore.

For example, we have begun working on a curriculum indicator data collection and reporting site to be available through the Internet. The goal is to provide a means for both electronic entry and reporting of curriculum indicator data for educators and researchers. Teachers using the system will be able to receive immediate feedback; a profile of their own practice (including a map of their instructional content); summary results of other teachers in their district, state, or nationally; and content maps for various assessment instruments. The site could be used in a number of ways that serve both research and professional development needs of the education community.

Using Video

We have a good deal of confidence in the instruments developed thus far, but we have no doubt that they could be improved. More work is needed in examining the validity and reliability of these instruments in order to ensure as accurate an indicator system as possible. Toward this end, we believe that work with video of classroom instruction holds tremendous potential. Video makes possible a tremendously flexible observation environment in which multiple observers can record descriptions of identical classroom lessons. Such analyses would undoubtedly provide a better understanding of how and why descriptions may vary and would likely lead to further improvements in the terminology and language used in data collection instruments.

Video lessons provide opportunities to examine issues of reliability and validity, and use of indicator instruments for describing lessons. In addition, video lessons provide a unique professional development opportunity for teachers to investigate varying forms of practice, to refine their language for describing differences in those practices, and to reflect upon the implications for their own instruction.

Extending Analyses and Use

We are also excited about a number of developments that will extend the types of possible analyses and the use of these instruments. For example, procedures are being developed to use the content taxonomies developed for mathematics and science in analyzing the content of curriculum standards, frameworks, and guidelines. This will provide additional measures of the *intended* curriculum in a metric that should allow careful comparison to the *enacted* and *assessed* curricula as described by instruments using a similar language or taxonomy.

The language and procedures we have developed for content analysis will allow for examination and description of other types of curricular documents as well. For example, instructional artifacts, such as assignments, classroom assessments, lab work, and portfolios provide yet another source for describing, analyzing, and comparing the *enacted* curriculum (Burstein et al., 1995). Using a consistent language to describe such artifacts will make it possible to check the validity of other data sources, such surveys and observations.

Finally, educators and professional development providers are beginning to turn to curriculum indicator data as an informational tool for teachers and schools to use in investigating their curriculum decisions. With funding from the National Science Foundation, we are currently using curriculum indicators in an experimental study to examine the effects of curriculum data on teacher practice when employed as a central component of a professional development package focused on data-driven decision-making. We have already found, less than a year into this study, that when teachers are presented with curriculum data and provided the opportunity to discuss the implications of the data, they become engaged and animated in the conversations. Whether such conversations lead to actual changes in practice is a key question that the study hopes to answer.

Summary and Conclusion

The past decade has seen growing interest in and improved quality of curriculum indicator data. Instruments for mathematics and science have undergone multiple revisions and field tests, new draft instruments for language arts and history have been developed, and the categories of cognitive demand have been carefully reworked. Numerous studies using our content taxonomies have been conducted and others studies are planned.

Of particular note has been the development of a systematic language for describing and comparing the *intended*, *enacted*, *assessed*, and *learned* curricula. This has facilitated the use of alignment analyses and led to preliminary results indicating the predictive validity of some alignment measures.

Growing in popularity among researchers, particularly evaluators of systemic reform, curriculum indicator data are also beginning to be used for school improvement, professional development, and teacher reflection. These broad and growing uses underscore the need for continued work in refining the language and instrumentation through investigation into their properties of reliability and validity. We see the use of video as making a valuable contribution to such investigations.

Other advances also appear on the horizon, such as the use of electronic data collection and reporting; content analyses of standards, frameworks, and guidelines; and opportunities for expanding the language and collaboration across research agendas. Each of these factors contributes to a sense of optimism that we are on the right track in pursuing a common and systematic language for describing key elements of the curriculum.

References

Aschbacher, P. R. (1999). *Developing indicators of classroom practice to monitor and support school reform*. Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing, University of California-Los Angeles.

Blank, R. K., Kim, J. J., and Smithson, J. L. (2000). *Survey results of urban school classroom practices in mathematics and science: 1999 Report*. Norwood, MA: Systemic Research, Inc.

Burstein, L., McDonnell, L. M., Van Winkle, J., Ormseth, T., Mirocha, J., and Guitton, G. (1995). *Validating national curriculum indicators*. Santa Monica, CA: RAND.

Council of Chief State School Officers. (2000). *Using data on enacted curriculum in mathematics and science: Sample results from a study of classroom practices and subject content.* Washington, DC: Author.

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

Kim, J. J., Crasco, L. M., Smithson, J. L. and Blank, R. K. (2001). *Survey results of urban school classroom practices in mathematics and science: 2000 report.* Norwood, MA: Systemic Research, Inc.

Mayer, D. P. (1999). Measuring instructional practice: Can policymakers trust survey data? *Educational Evaluation and Policy Analysis*, *21*(1), 29-45.

McKnight, C. C., Crosswhite, F. J., Dossey, J. A., Kifer, E., Swafford, J. O., Travers, K. J., and Cooney, T. J. (1987). *The underachieving curriculum: Assessing U.S. schools mathematics from an international perspective*. Champaign, IL: Stipes.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

National Research Council. (1996). *National science education standards*. Washington, DC: National Academy of Sciences, National Academy Press.

Porter, A. C. (1991). Creating a system of school process indicators. *Educational Evaluation and Policy Analysis*, *13*, (1), 13-29.

Porter, A. C. (1995). The uses and misuses of opportunity-to-learn standards. *Educational Researcher*, *24*(1), 21-27.

Porter, A. C. (1998a.) The effects of upgrading policies on high school mathematics and science. In D. Ravitch (Ed.), *Brookings papers on education policy* (pp. 123-172). Washington, DC: Brookings Institution Press.

Porter, A. C. (1998b). *Curriculum reform and measuring what is taught: Measuring the quality of education processes.* Paper presented at the annual meeting of the Association for Public Policy Analysis and Management, New York, NY.

Porter, A. C., Floden, R., Freeman, D., Schmidt, W., and Schwille, J. (1988). Content determinants in elementary school mathematics. In D. A. Grouws and T. J. Cooney (Eds.), *Perspectives on research on effective mathematical teaching* (pp. 96-

113). Hillsdale, NJ: Lawrence Erlbaum Associates. (Also Research Series 179, East Lansing: Michigan State University, Institute for Research on Teaching.)

Porter, A. C., Garet, M., Desimone, L., Yoon, K., and 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*. Washington, DC: U.S. Department of Education.

Porter, A. C., Kirst, M.W., Osthoff, E. J., Smithson, J. L., and Schneider, S. A. (1993). *Reform up close: An analysis of high school mathematics and science classrooms* (Final report to the National Science Foundation on Grant No. SAP-8953446 to the Consortium for Policy Research in Education). Madison, WI: University of Wisconsin-Madison, Consortium for Policy Research in Education.

Porter, A. C., and Smithson, J. L. (2001). Are content standards being implemented in the classroom? A methodology and some tentative answers. In S. H. Fuhrman (Ed.), *From the capitol to the classroom: Standards-based reform in the states* (pp. 60-80). Chicago: National Society for the Study of Education, University of Chicago Press.

Schmidt, W. H., McKnight, C. C., and Raizen, S. A. (with Jakwerth, P. M., Valverde, G. A., Wolfe, R. G., Britton, E. D., Bianchi, L. J., and Houang, R. T.). (1996). *A splintered vision: An investigation of U.S. science and mathematics education*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Schmidt, W. H., McKnight, C., Cogan, L., Jakwerth, P. and Houang, R. T. (1999). *Facing the consequences: Using TIMSS for a closer look at United States mathematics and science education.* Boston: Kluwer Academic Publishers.

Schwille, J. R., Porter, A. C., Belli, G., Floden, R., Freeman, D. J., Knappen, L. B., Kuhs, T. M., and Schmidt, W. H. (1983). Teachers as policy brokers in the content of elementary school mathematics. In L. Shulman and G. Sykes (Eds.), *Handbook on teaching and policy* (pp. 370-391). New York: Longman. (Also Research Series No. 113, East Lansing, MI: Michigan State University, Institute for Research on Teaching, 1983.)

Shavelson, R. and Stern, P. (1981). Research on teachers' pedagogical thoughts, judgments, decisions, and behavior. *Review of Educational Research*, *51*, 455-498.

Smithson, J. L. and Porter, A. C. (1994). *Measuring classroom practice: Lessons learned from efforts to describe the enacted curriculum — The reform up close study.* New Brunswick, NJ: Consortium for Policy Research in Education, Rutgers University.

Webb, N. L. (1997). *Criteria for alignment of expectations and assessments in mathematics and science education.* Madison, WI: Council of Chief State School Officers and National Institute for Science Education, University of Wisconsin.

Webb, N. L. (1999). *Alignment of science and mathematics standards and assessments in four states.* Madison, WI: Council of Chief State School Officers and National Institute for Science Education, University of Wisconsin.

Appendix A: Mathematics Topics

High School

Number Use/Operations
Estimation
Computational algorithms
Fractions
Decimals
Ratio and proportion
Percent
Real numbers
Number theory
Order of operations
Relationships between operations
Mathematical properties (e.g., the distributive property)

Consumer Applications
Simple interest
Compound interest
Rates (e.g. discount, commission)
Spreadsheets

Measurement
Use of measuring instruments
Theory (arbitrary, standard units, unit size
Conversions
Metric (SI) system
Length, perimeter
Area, volume
Surface area
Angles
Circles (pi, radius, diameter, area)
Pythagorean theorem
Mass (weight)
Time, temperature
Speed

Middle School

Number/sense, Properties, Relationships
Place value
Fractions
Decimals
Percent
Ratio, proportion
Integers
Real numbers
Exponents, scientific notation
Absolute value
Factors, multiples, divisibility
Odds, evens, primes, composites
Estimation
Order of operations
Relationships between operations
Mathematical properties (e.g., the distributive property)

Computation
Whole numbers
Fractions
Decimals
Percents
Ratio, proportion

Measurement
Use of measuring instruments
Theory (arbitrary, standard units, unit size)
Conversions
Metric (SI) system
Length, perimeter
Area, volume
Surface area
Direction, location, navigation
Angles
Circles (pi, radius, diameter, area)
Pythagorean theorem
Simple trigonometric ratios and solving right triangles
Mass (weight)
Time, temperature
Rates (including derived and direct)

Elementary School

Number/sense, Properties, Relationships
Place value
Patterns
Decimals
Percent
Real numbers
Exponents, scientific notation
Absolute value
Factors, multiples, divisibility
Odds, evens, primes, composites
Estimation
Order of operations
Relationships between operations

Operations
Add, subtract whole numbers
Multiplication of whole numbers
Division of whole numbers
Combinations of add, subtract, multiply and divide using whole numbers
Equivalent fractions
Add, subtract fractions
Multiply fractions
Divide fractions
Combinations of add, subtract, multiply and divide using fractions
Ratio, proportion
Representations of fractions
Decimal equivalent to fractions
Add, subtract decimals
Multiply decimals
Divide decimals
Combinations of add, subtract, multiply, and divide using decimals

Measurement
Use of measuring instruments
Units of measure
Conversions
Metric (SI) system
Length, perimeter
Area, volume
Surface area
Telling time
Circles (e.g. pi, radius, area)
Mass (weight)
Time, temperature

High School (cont.)

Algebraic Concepts
Integers
Absolute value
Exponents, scientific notation
Use of variables
Expressions
Evaluation of formulas & expressions
One-step equations
Coordinate plane
Multi-step equations
Inequalities
Literal equations
Lines/slope and intercept
Operations on polynomials
Factoring
Square root and radicals
Operations on radicals
Rational expressions

Advanced Algebra
Quadratic equations
Systems of equations
Systems of inequalities
Compound inequalities
Matrices/determinants
Conic sections
Rational, negative exponents/ radicals
Rules for exponents
Complex numbers
Binomial theorem
Factor/remainder theorems
Field properties of real number systems

Data Analysis
Bar graph, histogram
Pictographs
Line graphs
Stem and leaf plots
Scatter plots
Box plots
Mean, median, mode
Mean deviation
Smoothing of graphs

Middle School (cont.)

Algebraic Concepts
Absolute value
Use of variables
Evaluation of formulas & expressions
One-step equations
Coordinate plane
Multi-step equations
Inequalities
Linear, non-linear relations
Operations on polynomials
Factoring
Square roots and radicals
Operations on radicals
Rational expressions
Functions and relations
Quadratic equations
Systems of equations
Systems of inequalities
Matrices/determinants
Complex numbers

Data Analysis/Probability/ Statistics
Bar graph, histogram
Pie charts, circle graphs
Pictographs
Line graphs
Stem and leaf plots
Scatter plots
Box plots
Mean, median, mode
Line of best fit
Quartiles, percentiles
Sampling, sample spaces
Simple probability
Compound probability
Combinations and permutations
Summarize data in a table or graph

Elementary School (cont.)

Algebraic Concepts
Expressions, number sentences
Equations (e.g., missing value)
Absolute value
Function (e.g., input/output)
Integers
Use of variables, unknowns
Inequalities
Properties
Patterns

Probability and Statistics
Bar graph, histogram
Pictographs
Line graphs
Mean, median, mode
Quartiles, percentiles
Simple probability
Combinations and permutations
Summarize data in table or graph

High School (cont.)

Functions
Notation
Relations
Linear
Quadratic
Polynomial
Rational
Logarithmic
Exponential
Trigonometric/circular
Inverse
Composition

Geometric Concepts
Basic terminology
Relationships between lines & their parts, angles, and planes
Triangles
Quadrilaterals
Polygons
Congruence
Similarity
Parallels
Circles
Constructions

Advanced Geometry
Logic, reasoning, proof
Symmetries
Loci
Spheres, cones, cylinders
Polyhedra
3-dimensional relationships
Transformational
Coordinate
Vectors
Analytic
non-Euclidean
Topology

Trigonometry
Basic ratios
Radian measure
Right triangle trigonometry
Law of sines, cosines
Identities
Trigonometric equations
Polar coordinates
Periodicity
Amplitude

Middle School (cont.)

Geometric Concepts
Basic terminology
Points, lines, rays, and vectors
Patterns
Congruence
Similarity
Triangles
Quadrilaterals
Circles
Angles
Polygons
Polyhedra
Models
Symmetry
Spatial reasoning, 3-D relationships
Transformations (e.g., flip, turn)
Pythagorean theorem
Simple trigonometric ratios

Elementary School (cont.)

Geometric Concepts
Basic terminology
Points, lines, rays, and vectors
Patterns
Congruence
Similarity
Triangles
Quadrilaterals
Circles
Polygons
Polyhedra
Symmetry
Models
Spatial reasoning, 3-D relationships
Transformations (e.g., flip, turn)

High School (cont.)

Statistics
Variability, standard deviation
Quartiles, percentiles
Bivariate distributions
Sampling
Confluence intervals
Correlation
Lines of best fit
Hypothesis testing
Chi-square
Data transformation
Central limit theorem

Probability
Sample spaces
Compound probability
Conditional probability
Independent/dependent events
Empirical probability
Expected value
Binomial distribution
Normal curve

Finite Math/Special Topics
Sets
Logic
Mathematical induction
Linear programming
Networks
Iteration/recursion
Permutations, combinations
Simulations
Fractals

Analysis
Sequence and series
Limits
Continuity
Rates of change
Maxima, minima
Differentiation
Integration

Technology
Use of calculators or computers
Use of the internet
Computer programming

Middle School (cont.)

Technology
Use of calculators
Graphing calculators
Computers and the internet

Elementary School (cont.)

Technology
Use of calculators
Computers and the internet

Appendix B: Science Topics

High School

Nature of Science
Nature and structure of science
Nature of scientific inquiry
History of science
Ethical issues/Critiques of science
Science, technology, & society

Measurement & Calculation in Science
The international system
Mass & weight
Length
Volume
Time
Temperature
Accuracy & precision
Significant digits
Derived units
Conversion factors
Density

Components of Living Systems
Cell structure/function
Cell theory
Transport of cellular material
Cell metabolism
Cell response
Genes
Cell specialization

Biochemistry
Living elements (C,H,O,N,P)
Atomic structure & bonding
Synthesis reactions (Proteins)
Hydrolysis
Organic compounds: Carbon, Proteins, Nucleic/Amino Acids/ Enzymes

Botany
Nutrition/Photosynthesis
Circulation
Respiration
Growth/development/behavior
Health & disease

Middle School

Nature of Science
Scientific habits of mind (e.g., reasoning, evidence-based conclusions, skepticism)
Scientific method (e.g., observation, experimentation, analysis, theory development, and reporting)
History of scientific innovations
Ethical issues in science

Measurement & Calculation in Science
The international system
Mass & weight
Length
Volume
Time
Temperature
Accuracy & precision
Significant digits
Derived units
Conversion factors
Density

Science, Health, & Environment
Personal health, behavior, disease, nutrition
Environment health, pollution, waste disposal
Resources, conservation
Natural and human caused hazards

Components of Living Systems
Cell structure/function
Cell theory
Cell response
Genes
Organs
Organ systems

Botany
Nutrition/Photosynthesis
Vascular system
Growth/development/behavior
Health & disease

Elementary School

Nature of Science
Nature and structure of science
Nature of scientific inquiry
History of science
Ethical issues/Critiques of science
Science, technology, & society

Measurement & Calculation in Science
The international system
Mass & weight
Length
Volume
Time
Temperature
Density

Components of Living Systems
Structure & function in plants
Structure & function in animals

Botany
Nutrition/Photosynthesis
Reproduction
Growth/development/behavior

High School (cont.)

Animal Biology
Nutrition
Circulation
Excretion
Respiration
Growth/development/behavior
Health & disease
Skeletal & muscular system
Nervous & endocrine system

Human Biology
Nutrition/Digestive system
Circulatory system (Blood)
Excretory system
Respiration & respiratory system
Growth/development/behavior
Health & disease
Skeletal & muscular system
Nervous & endocrine system

Genetics
Mendelian genetics
Modern genetics
Inherited diseases
Biotechnology
Human genetics

Evolution
Evidence for evolution
Lamarckian theories
Modern evolutionary theory
Life origin theories
Natural selection
Classification
Adaptation & variation

Reproduction & Development
Mitotic/Meiotic cell division
Asexual reproduction
Sexual reproduction & development in plants
Sexual reproduction & development in animals
Sexual reproduction & development in humans

Middle School (cont.)

Animal Biology
Nutrition
Circulation
Excretion
Respiration
Growth/development/behavior
Health & disease
Skeletal & muscular system
Nervous & endocrine system

Human Biology
Nutrition/Digestive system
Circulatory system (Blood)
Excretory system
Respiration & respiratory system
Growth/development/behavior
Health & disease
Skeletal & muscular system
Nervous & endocrine system

Evolution
Evidence for evolution
Modern evolutionary theory
Human evolution
Classification
Natural selection
Adaptation & variation

Reproduction & Development
Mitotic/Meiotic cell division
Asexual reproduction
Inherited traits
Sexual reproduction & development in plants
Sexual reproduction & development in animals
Sexual reproduction & development in humans

Elementary School (cont.)

Animal Biology
Nutrition
Respiration
Growth/development/behavior

Human Biology
Nutrition/Digestive system
Body systems
Respiration

Growth, development, & behavior
Reproduction & development
Life cycles in plants
Life cycles in animals

High School (cont.)

Ecology
Nutritional relationships
Competition & cooperation
Energy flow relationships
Ecological succession
Ecosystems
Population dynamics
Environmental chemistry

Energy
Potential energy
Kinetic energy
Conservation of energy
Heat energy
Light energy
Sound energy
Thermal expansion & transfer
Work & energy
Nuclear energy

Motion & Forces
Vector & scalar quantities
Displacement as a vector quantity
Velocity as a vector quantity
Relative position & velocity
Acceleration
Newton's First Law
Newton's Second Law
Newton's Third Law
Momentum, impulse, and conservation
Equilibrium
Friction
Universal gravitation

Electricity
Static electricity: production, transfer, & distribution
Coulomb's law
Electric fields
Current electricity
Current, voltage, & resistance
Series & parallel circuits
Magnetism
Effects of interacting fields

Middle School (cont.)

Ecology
Food chains/Webs
Competition & cooperation
Energy flow relationships
Ecological succession
Ecosystems
Population dynamics

Energy
Potential energy
Kinetic energy
Work & force
Conservation of energy
Heat energy
Mechanical energy & machines
Nuclear energy

Motion & Forces
Velocity
Mass
Newton's First Law
Newton's Second Law
Newton's Third Law
Forces
Friction
Universal gravitation

Science & Technology
Design a solution or product, implement a design
Relationship between scientific inquiry and technological design
Technological benefits, trade-offs, and consequences

Electricity
Static electricity: production, transfer, & distribution
Coulomb's law
Electric fields
Current electricity
Series & parallel circuits
Magnetism

Elementary School (cont.)

Ecology
Food chains/Webs
Ecosystems - Change/Impacts
Renewable resources
Pollution & conservation
Human population growth

Energy
Forms of energy
Conservation of energy
Transfer of energy
Motion & forces
Position
Speed
Forces

Electricity
Current electricity
Series & parallel circuits
Magnetism

High School (cont.)

Waves
Characteristics and behavior
Light
Electromagnetic
Sound

Kinetics & Equilibrium
Molecular motion
Pressure
Kinetics and temperature
Equilibrium
Reaction Rates

Properties of Matter
Characteristics & composition
States of matter (S-L-G)
Physical & chemical changes
Physical & chemical properties
Isotopes, atomic number, & atomic mass
Atomic theory
Quantum theory & Electron clouds

Earth Systems
Earth's shape, dimension and composition
Earth's origins and history
Maps, locations and scales
Measuring using relative and absolute time
Mineral & rock formations & types
Erosion & weathering
Plate tectonics
Formation of: volcanoes, earthquakes, & mountains
Evidence of change
Dynamics & energy transfer
Oceanography

Middle School (cont.)

Characteristics & behavior of Waves
Light
Electromagnetic
Sound

Kinetics
Molecular motion
Pressure
Kinetics and temperature

Properties of Matter
Characteristics & composition
States of matter (S-L-G)
Physical & chemical changes
Physical & chemical properties
Isotopes, atomic number, & atomic mass
Atomic theory

Earth Systems
Earth's shape, dimension and composition
Earth's origins and history
Maps, locations and scales
Measuring using relative and absolute time
Mineral & rock formations & types
Erosion & weathering
Plate tectonics
Formation of: volcanoes, earthquakes, & mountains
Oceanography

Elementary School (cont.)

Characteristics & behavior of Waves
Light
Sound

Properties of Matter
Characteristics & composition
States of matter (S-L-G)
Physical changes
Physical properties

Earth Systems
Earth's shape, dimensions, & composition
Soil composition
Surface characteristics
Evidence of change
Erosion & weathering

High School (cont.)

Astronomy
Stars
Galaxies
The solar system
Earth's moon
Earth's motion (rotation & revolution)
Location, navigation, & time

Meteorology
The Earth's atmosphere
Air pressure & winds
Evaporation/condensation/precipitation
Weather
Climate

Elements & The Periodic System
Early classification system
Modern periodic table
Interaction of elements
Element families & periods

Chemical Formulas & Reactions
Names, symbols, & formulas
Molecular & empirical formulas
Representing chemical changes
Balancing chemical equations
Stoichiometric Relationships
Oxidation/Reduction reactions
Chemical bonds
Electrochemistry
The Mole

Acids, Bases, & Salts
Arrhenius, Bronsted-Lowry, & Lewis Theories
Naming acids
Acid-Base behavior/strengths
Salts
pH
Hydrolysis
Buffers
Indicators
Titration

Middle School (cont.)

Astronomy
Stars
Galaxies
Asteroids and comets
The solar system
The Earth's moon
The Earth's motion: rotation & evolution
Location, navigation, & time

Meteorology
The Earth's atmosphere
Air pressure & winds
Evaporation/condensation/precipitation
Weather
Climate

Elements & The Periodic System
Early classification system
Modern periodic table
Interaction of elements
Characteristics of elements

Chemical Formulas & Reactions
Names, symbols, & formulas
Molecular formulas
Representing chemical change
Chemical bonds
Types of reactions

Acids, Bases, & Salts
Naming acids
Acid-Base behavior/strengths
Salts
pH
Hydrolysis
Indicators

Elementary School (cont.)

Astronomy
Stars
Galaxies
The solar system
The Earth's moon
The Earth's motion: rotation & revolution
Location, navigation, & time

Meteorology
The Earth's atmosphere
Air pressure & winds
Evaporation/condensation/precipitation
Weather
Climate

High School (cont.)

Organic Chemistry
Hydrocarbons, Alkenes, Alkanes, & Alkynes
Aromatic Hydrocarbons
Isomers & Polymers
Aldehydes, Ethers, Ketones, Esters, Alcohols, Organic Acids
Organic Reactions
Carbohydrates, Proteins, Lipids

Nuclear Chemistry
Nuclear structure
Nuclear equations
Fission
Radioactivity
Half-life
Fusion

Middle School (cont.)

Environmental Chemistry
Pollution
Acid rain
Ozone depletion
Toxic & nuclear waste
Greenhouse effect

Nuclear Chemistry
Nuclear structure
Fission
Radioactivity
Fusion

Elementary School (cont.)

Appendix C: Mathematics Cognitive Demand

Memorize Facts, Definitions, Formulas

Classroom activities focused on this level of cognitive demand include recall of traditional math skills and knowledge, e.g., recall of definitions, technical skills such as factoring polynomials, standard algorithms, basic number facts, and operations. In activities focused on memorization, students spend much time learning (memorizing) traditional computational procedures. Such activities focus on basic skills and paper and pencil computation.

Students:

Recall basic geometric terminology.
Recall the formula for the area of a circle.
Recite multiplication facts.
Tell the formula for finding percent.
Name a right angle in a diagram.

In grade 4, students will memorize number facts for the four basic operations.

In grade 7, students will memorize the different kinds of angles.

Communicate Understanding of Concepts

- Communicate mathematical ideas.
- Use representations to model mathematical ideas.
- Explain findings and results from analysis of data.

At this level of cognitive demand, students share their mathematical understandings in both oral and written form with their teacher and classmates. Students actively participate in conversations about mathematics. They talk to other students about mathematics (e.g., critique, question). If a student gives an incorrect response, the teacher may discuss the incorrect response with the student inviting other students to participate. The following is an example of a conceptual approach to understanding percent taken from *Mathematics in Context* (van den Heuvel-Panhuizen et al., 1997): Two shop keepers are comparing their prices. Barbara's store sells a watch for $20. Dennis's store sells the same watch for $40. Barbara says, "Your store price is 100 percent more expensive!" "That's not true," says Dennis. "Your store price is only 50 percent less." Who is right?

Students:

- Generate and describe number sequences involving constant multiplication and division or combinations of operations.
- Select the relevant information to solve a problem and determine what additional information is needed.
- Show that the operation of multiplication is the inverse of division.
- Describe two features of a decimal number.
- Explain their strategy to others.

In grade 4, students explain what makes a geometric shape a triangle.

In grade 7, students use mathematical language and symbols to represent problem situations.

Perform Procedures/Solve Routine Problems

- Do computations.
- Make observations.
- Take measurements.
- Solve routine story problems.

In classroom activities focused on this level of cognitive demand, the emphasis is on the product (e.g., answer) rather than the process (e.g., strategy). This aspect of math is concerned with getting procedural answers to particular questions. Students demonstrate fluency with basic skills by using these skills accurately and automatically, and demonstrate practical competence with other skills by using them effectively to accomplish a task. In activities focused on performing procedures and solving routine problems, students may be asked to select and apply various computational methods, including mental math, paper and pencil techniques, and the use of calculators. The following is an example of a routine problem (assuming that students already know the algorithm): Sam has two cards. Diane has three cards. How many do they have altogether?

Students:

- Use standard algorithms to solve computational problems.
- Evaluate formulas using both pencil and paper and more advanced technology.
- Solve equations symbolically.
- Use standard methods to solve basic problems.
- Find the area of a triangle.
- Solve $3x + 4 = 13$.
- Divide fractions.

In grade 4, students will use the four basic arithmetic operations in a variety of computational problems.

In grade 7, students will use a formula to find the percent of a number.

Solve Non-routine Problems/Make Connections

- Apply and adapt a variety of appropriate strategies to solve non-routine problems.
- Apply mathematics in contexts outside of mathematics.
- Analyze data, recognize patterns.

In activities focused on this level of cognitive demand, students apply their math knowledge creatively to solve problems in unfamiliar problems. Many multi-step problems fall into this category. Non-routine problems may be open-ended problems with more than one right answer or problems where the answer is not obvious if the student follows a standard step-by-step routine. Non-routine problems often may be solved in more than one way.

Making connections means that students see relationships between different topics and draw on these relationships in future mathematical activity. This applies within mathematics (e.g., relationships between algebra and geometry), and to other content areas (e.g., use of mathematics in science). The following non-routine problem taken from *Mathematics in Context* (van den Heuvel-Panhuizen et al., 1997) requires students to use previous work with percents to make connections:

The government of Elbonia is having problems accounting for all of the money spent. Mr. Butler is the Elbonian bureaucrat whose job is to deliver the money to developing countries. For his work, he gets a one percent commission. An undercover detective who is interviewing all the bureaucrats succeeds in getting a dinner appointment with Mr. Butler. After dinner, the server brings the check to the table. The total is $20. Mr. Butler announces his intention to leave a 15 percent tip. First, he gives the server a dollar. "That's five percent," he says. Then, he adds a dime to the dollar. "This is another 10 percent, so altogether it is a 15 percent tip," he explains. Suddenly, the detective jumps up and says, "Aha! Now I know where the money went! You are under arrest!"

1. What did the detective figure out that could be used to convict Mr. Butler of fraud? Include all the important information you know about percents so that the prosecuting attorney can convince the jury.

2. Is there any way that Mr. Butler could plan his defense? Explain.

Students:

- Work on problems for which there is no immediately obvious method of solution.
- Explain and support their solution strategy.
- Explain the connection between the greatest common factor of two numbers and their common multiple.

In grade 4, students recognize role of mathematics in their daily lives.

In grade 7, students apply mathematical problem solving to other content areas (e.g., measurement in science).

Conjecture/Generalize/Prove

- Complete proofs.
- Make and investigate mathematical conjectures.
- Find a mathematical rule to generate a pattern or number sequence.
- Determine the truth of a mathematical pattern or proposition.

In activities focused on this level of cognitive demand, students are making and justifying conjectures, not just learning techniques. Proof is a central concept in mathematics. It is important because based on explicit hypotheses, a proof shows that certain consequences follow logically, and these logical consequences can be used to build mathematical theories. There are several kinds of proof:

1. **Enactive proof:** Enactive proof involves carrying out a physical action to demonstrate the truth of something. It involves physical movement to show a relationship. A typical enactive proof is to demonstrate that a triangle with equal sides has equal angles by cutting out a triangle and folding it down its axis of symmetry to show that when the sides match so do the base angles.

2. **Visual proof:** A visual proof may involve enactive elements but usually has verbal or written support. A classic visual proof is the famous Indian proof of Pythagoras where four copies of a right triangle are placed in two different ways in a square.

3. **Manipulative proof:** Manipulative proof is often seen in algebra. For example, to show that (a + b)(a-b) = a2- b2, students multiply out the brackets on the left hand side and cancel the terms ba and -ab.

4. **Euclidean proof:** This is the classic formal proof of definitions, axioms and theorems. Mathematical proving consists of thinking in a logical manner, formulating and testing conjectures, and formulating and justifying statements, inferences, and conclusions. The following is an example problem that requires students to complete a proof:

Put these statements in order, and complete if necessary, so that they constitute a proof:

Two even numbers add to make an even number; if I divide an even number by two, there is no remainder; if I divide a number by two, it either goes exactly or there is a remainder of one; an even number can be written as: 2n;2n+2m=2(m+n).

Students:

- Develop and support mathematical conjectures.
- Justify why a rule works. (e.g., An odd number plus an odd number equals even number because if you take away from the odd number it will be even, so if you add the two numbers left over together, that makes an even number and three evens make an odd number.)
- Correct an argument. (e.g., $2n + 2n + 1 = 4n + 1$, which is odd. So even + even is odd.)
- Demonstrate that the product of two odd numbers is always odd.

In grade 4, students justify their answers and solution process in a variety of problems.

In grade 7, students follow and construct logical arguments and judge their validity.

NOTICE

Reproduction Basis

☐ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☒ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").