

HOW DOES CURRICULUM AFFECT LEARNING?

SCHOOLS MATTER. This statement is a truism to most. However, it must be followed by a statement of why schools matter, especially in light of the current debate surrounding the quality of American public education. Working from the seemingly simple belief that schools matter, assumptions will be made and policies will follow them on how to improve the quality of public schools. If there is no statement of why schools matter based on empirical data, the assumptions may be wrong and the policies may hinder rather than help improve the quality of public education.

Some have argued that the key to quality education lies with policies concerning students and their characteristics. These policies lead to practices such as grouping students into tracks so the “right” students get the “right” opportunities. This type of practice is based on the belief that socioeconomic status, parents’ education, student aptitude, and other background factors are more important to students’ achievement than what happens in schools.¹ In short, according to this belief schools matter only if the right students are in the right classes at the right time, and this is true no matter how effective those classes may be.

We believe that schools do matter; we have written this book to try to show why they matter using data on curriculum and achievement from a cross-national study—the Third International Mathematics and Science Study (TIMSS). Schools matter in many ways—they distribute resources, they create learning climates, they provide opportunities for learning experiences, and so on. This book focuses on one aspect of schooling: curriculum. Curriculum represents the intended courses of study and sequences of learning opportunities in formal schooling. We believe the TIMSS data clearly show that curriculum affects learning.

How does curriculum affect learning? Surely the answers to this question are essential to meaningful educational changes, yet the question is far from simple. To even grasp the question we need an understanding of curriculum, of learning, and of ways in which the two may interact. This question helps provide a *raison d'être* for large-scale cross-national comparative studies of achievement. Without such serious questions, those studies become nothing more than exercises in politics. The authors believe that the most important purpose for these studies, and for the TIMSS in particular, is to answer such questions. It is to those answers that this book is devoted.

The Question of Curriculum

Taking the question of how curriculum affects learning as our starting point, we need to investigate what we mean by curriculum and by learning, and how these interact. We begin first with the question of what we mean by curriculum.

The term *curriculum* comes from words meaning to “run a (race) course” and refers to a sequence of steps or stages in teaching and learning specific content. If we think of curriculum as a sequence of learning experiences, we immediately run into the difficulty that no one—teacher or otherwise—can consistently control the experiences of individual students. All that can be done is to provide students with opportunities to learn specific content. Thus, a good definition for curriculum is a sequence of learning opportunities provided to students in their study of specific content.

There is one obvious difficulty with curriculum being defined as sequences of learning opportunities. Such curriculum is invisible. We cannot see “sequences of learning opportunities.” We can see the plans for such sequences. We can see classroom activities meant to serve as opportunities to learn specific content in the sequence. We can see textbook pages that help provide those learning opportunities. However, we cannot see curriculum directly. We can see the artifacts and effects of curriculum, but not curriculum itself.

As a sequence of learning opportunities, curriculum has several aspects. It exists as plans and intentions—the sequence of learning opportunities that one wishes students to experience. It exists as patterns of classroom activities that are meant to implement those plans and provide the desired learning opportunities. It exists in textbooks as pages intended to support or present those classroom learning opportunities. It leaves marks on what time is devoted to specific contents by teachers or textbooks. It impacts

what students obtain as a result of opportunities to learn. Curriculum has these many different aspects and indications.

When we wished to study curriculum as a part of TIMSS, one of our first tasks was to choose artifacts and effects of curriculum we thought would reflect the various aspects of curriculum—intentions, implementations, and attainments. The latter was represented by student achievement on the TIMSS tests designed to measure student attainments at particular ages or grades in specific aspects of mathematics and the sciences. Assessing intentions and implementations was more problematic.

What artifacts (documents, books, lesson plans, etc.) should be taken as indicators of the intentions and implementations of curriculum? What aspects of classroom activities should be taken as indicators of curriculum by its effects? Although several artifacts were examined, we focus on four: content standards, textbooks, teachers' content goals, and duration of content coverage.

First, official documents often provide direct statements of the content and performance levels desired for students. Let us call these content standards. We systematically collected content standard documents from the countries participating in TIMSS. We then analyzed the specific science and mathematics content specified in the documents and the kinds of performance abilities expected from students. These content standards documents were taken as indications of curriculum as intention—as plans and goals.

Second, student textbooks are used in virtually all countries to support or more directly guide the learning opportunities of children. Textbooks represent ways in which curriculum can be implemented as opportunities to learn. We systematically collected representative samples of student textbooks in all participating countries for TIMSS Populations 1 and 2 (essentially fourth and eighth grade) and for the specialized content (e.g., calculus and advanced mathematics, and physics) of Population 3 (the end of secondary school). We have analyzed many aspects of those textbooks and continue to analyze others. For the purposes of this book, we focus primarily on the proportion of textbook space devoted to specific content areas in mathematics and the sciences. We consider this indicator of curriculum to be a bridge, expressing both curriculum as intention and, potentially, curriculum as implementation, depending on how and if specific parts of the analyzed textbooks are actually used in classroom instruction.

Finally, we wished to have some indicator of curriculum as it is actually implemented by teachers in attempting to provide learning opportunities for their students. Of the range of possibilities explored, two such

indicators are used in this book. We consider teachers' indicated learning goals and time coverage as our third and fourth indicators of curriculum. We consider these effects to be indicators of curriculum as implemented. They can be analyzed for the proportion of time or emphasis devoted to various contents.

Thus, we use these four aspects of curriculum—content standards, text-book space, teacher content goals, and duration of content coverage—to make visible the invisible. We use data on these four factors to search for answers to how curriculum affects learning. These were not the only choices possible or the only that we explored. They are, however, sufficient to reveal some important things about how curriculum matters, as will be seen later in this book.

The Question of Learning

To answer the question of how curriculum affects learning, we must not only have some understanding of curriculum; we must also have some understanding of learning.

When we talk about learning we are not talking about understanding the cognitive mechanisms of individual learning, which vary among children and among cultures. We are talking about gains in competencies and knowledge, about growth in attainment, and about what happens in schools. We are interested in what affects gains in the achievements of children. The empirical data with which we work are the results of the TIMSS achievement tests.

We put safeguards in place in our collection of TIMSS data to ensure that the samples are representative of each country's children at a particular age or grade.² We also made every reasonable effort to find high-quality test items that measure relevant mathematics and science contents and student competencies so that the resulting performances are typical of what those children can do repeatedly, and accurately reflect the nature and level of those children's learning.³ Neither process was perfect but both were reasonably successful.⁴ Now that the data have been collected, equally careful efforts are needed for scale development, analysis, interpretation, and reporting.

In official international reports of TIMSS achievement data, most scores were reported for large collections of items. Aggregate scores were reported for mathematics and for science at each sampled population. These scores were based on the entire collection of mathematics and science items, respectively, at each test level. Slightly more specific scores were reported for broad categories within mathematics and within the sci-

ences (algebra, earth sciences, etc.). Attention was focused on the comparative status of the achievement of each participating country's students.

Does comparative achievement status using such broad categories reflect learning? It does so only in a broad, cumulative sense—what has been learned in broad content areas over many years of schooling. Such broad measures are unlikely to be sensitive to the specifics of curriculum coverage in mathematics and the sciences, especially at any single grade level. Learning consists of change and gain in educational attainments. Curriculum specifics are most relevant to such gains.

TIMSS focused on three populations. Population 1 consisted of the two adjacent grades containing the majority of nine-year-olds in each country. Population 2 consisted of the two adjacent grades containing the majority of thirteen-year-olds in each country. Population 3 consisted of all students in the last year of secondary school with subpopulations for those still studying advanced mathematics or physics or both. The first two populations allowed for a focus on students of a specific age or in a specific grade.

Tests were given toward the end of the school year in each case. By drawing samples carefully from two adjacent grades, it was possible to measure achievements in both grades and to construct an indication of gain from one grade's school experience at the national level. This was not a truly longitudinal study because the same students were not followed throughout a period of time. No gain data could be reported for individual students. However, the TIMSS data might be considered "quasi-longitudinal." By measuring similar students in each country at two close periods of time separated mainly by one year's instruction, gains seen in the higher grade could be interpreted as an indication of what was learned in mathematics and the sciences during that year's instruction. Although these gain data could be estimated only in the aggregate such as at the national level, they still served as a better indicator of learning in a specific grade than comparative achievement status. The latter at best indicated something about cumulative learning.

Measuring learning in ways that are sensitive to curriculum factors is enhanced not only by using gain rather than status scores, but also by using measures created by combining only those items more specific to particular topic areas. Common practice is to combine the items into a total score or some other scale value related to total score (usually using item response theory such as Rasch scaling). A common feature of this approach is an attempt to measure a single trait or competency that lies behind success in all of the items combined. Answering more items correctly in a set to be scaled is taken to imply possession of more of the trait

or a greater amount of the underlying competency. Unfortunately, a set of items measuring diverse topic areas yields only a measure of whatever common competency lies behind all those items. If the topics represented by the items are very diverse, the underlying competency must be more general.

This practice of combining items into a total score (or a similarly broad category score such as physics or algebra) creates a crucial trade-off. More items permit a more accurate estimate of an underlying competency. However, in cross-national comparative studies of student achievement, there is always a desire to cover a broad range of content with as few items as possible given limited testing time. As a result, even when mathematics or science items are grouped into broad categories (algebra, earth science, etc.), the content of these categories is very diverse. When scales are for all of mathematics or all of science at a specific grade level, the sets of items are diverse indeed and the competencies measured correspondingly are even more general.

Why is this a problem in studying the relationship between curriculum and learning? Put simply, the more general the competency measured, the more likely it is to be influenced by factors outside of schooling such as motivation, social class, and general aptitude. An achievement measure that is sensitive to curriculum differences must draw on one or more specific competencies that are affected by learning opportunities provided by those curricula. Differences among curricula should show differences in patterns of achievement for large samples of students—for example, in national estimates. Diverse sets of items measuring only very general competencies are far less likely to yield measures that are specific enough to be affected by curriculum differences in a given year. For the purposes of studying the effect of curriculum on learning, greater diversity and the resulting measures of more general competencies are less desirable. Use of more specific, closely related sets of test items yield measurements of more specific skills relevant to those items.

Thus, closely related item sets are more likely to be sensitive to differences among curricula in how content needed for correct responses to those closely related items are treated in providing students with an opportunity to learn. Unfortunately, limitations on testing time and the need for broad content coverage make it virtually impossible to include enough sets of closely related items to provide scale scores that measure all the things one would like to compare among students of similar ages or grades from different countries. In studying curriculum and learning, however, a focus on smaller, more closely related item sets may yield measures of learning that are more sensitive to curriculum differences.

Curriculum, Learning, and Culture

In first addressing the question of how curriculum affects learning, we said we needed to know not only something about learning and curriculum, but also about how the two are related. Much of that relationship is discussed in the subsequent chapters of this book. First, there is one aspect of the link between learning and curriculum that we wish to discuss briefly here.

We believe that there is a close relationship between curriculum, learning, and culture. The term *culture* can have many different meanings—everything from “high” culture (“the best that has been thought and felt” in a country or civilization, according to Matthew Arnold) to the ideology of a particular subgroup within a society (their conceptual and political way of interpreting their world). We want to focus on “lived culture”—how everyday life for children in school is experienced and shaped through the social institutions in their country. We wish to examine the relationship between curriculum and learning for each country that participated in TIMSS, so we focus on “lived culture” at a national level.

Explorations of culture are not easily approached through quantitative data. True understanding requires a wealth of qualitative data and investigation. That is beyond the scope of the TIMSS data on curriculum and achievement. Thus, what we offer here is not what can be proved about the relation of curriculum, learning, and culture, but rather a hypothesis that seems revealing and appropriate given the relationships we have seen in the quantitative data on curriculum and achievement.

We will offer a variety of cultural hypotheses throughout this report, but each is a variation of a more general one we might call *the* cultural hypothesis: how curriculum matters to learning is affected by how curriculum is shaped by the lived culture that affects schooling and school experiences.

Although it is an oversimplification, we may say that the past shapes our schools and our schools shape the future. A nation’s culture or cultures shape its history and self-image, including values, institutions, goals, and the events that unfold from these. A nation’s educational system is shaped by these same factors—even the very idea of what schools are and should accomplish. Curricula and other policies express cultural values, goals, and commonly accepted ways to reach those goals. The same forces give particular form to educational systems.

The most fundamental question is not *whether* culture has an impact on learning, but rather *how* culture has an impact on learning. Culture acts in both direct and indirect ways. Culture shapes what we value and

thus enters our judgments in everyday living. Culture also acts less directly by shaping the social institutions we encounter, which in turn help create our everyday experiences. In this context, institutions such as school, grade level, subject area, and others emerge. They are shaped by the socialization through which education's institutions are formed and change.

Children encounter formal schooling primarily through education's social institutions. The organization of education within a culture is shaped to determine who receives which opportunities and when they receive them. This does not imply that the informal learning of socialization stops when formal schooling begins, but rather that the informal and the formal proceed side by side throughout the years of schooling.

The more diffuse impacts of culture are left for ethnographic study and cultural analysis—we do not attempt to address them here. They are certainly real, but they are beyond the scope of even the complex TIMSS data. On the other hand, insights into how social institutions function and, in particular, how education's social institutions function are accessible through quantitative data. The TIMSS data seems to offer strong possibilities here. We want to view how culture functions indirectly through education's social institutions.

Within a country, socioeconomic status, race, gender, and subculture memberships can have the same sort of impacts as those of differing countries in cross-national comparisons. These factors have indirect as well as direct effects. In the United States and some other countries, they make an impact through local control. Local control sets up local school subcultures that have differing impacts on curriculum, which, in turn, impacts achievement. The lack of an institutional center for curriculum in the United States (see Chapter Four) allows for greater differences in curriculum and different impacts on learning. For the most part this book focuses on differences between countries, although there are some ventures into intra-country analyses in Chapter Ten.

We will focus on selected aspects of educational institutions—educational decision making, subject matter content, and so forth—and how these vary among national education systems formed through different cultural and institutional histories. *Institution* here refers not to formal organizations or buildings but rather to ways of conducting and organizing aspects of education, to what sociologists mean when they talk about social institutions and institutionalized aspects of society. We hope to demonstrate the clear impact of those differences in educational institutions (in the sociological sense). We hope the understanding of these differences will make clear relationships in the data of curriculum and

achievement that would remain unclear if these data were approached without cultural awareness or with more simplistic views of culture and social institutions.

For the present analyses, we will focus on four key aspects of social educational institutions. First, we recognize the goals and purposes of education in society as an institution not because we study it but because of its centrality from a conceptual point of view. Education's goals and purposes reflect cultural beliefs and values. These goals and purposes are institutionalized in different ways in different nations and education systems. This affects many aspects of how goals and intentions shape education practically. For example, in some cultures and countries, official goals remain at very global levels, while in others they are specific and myriad. This "grain size" of curricular goals seems likely to have an impact on how formal education is conducted. We need to explore the differences among these national methods of articulating educational goals and purposes because they both affect and are affected by that which comes afterward in education. For the most part, however, this is a study that must be left to others. It was not what TIMSS was designed to investigate.

Two key aspects of schooling affected by educational goals and purposes are authority within educational institutions and how authority is exercised in practical decision making. Surveying the arena of goals and authority cross-nationally makes it clear that specific goals are not wedded to specific arrangements for the distribution of authority and decision making. The centers of decision-making authority vary greatly among countries whether at the national, local, or school levels. For example, Switzerland has no school principals at certain grade levels, so does not even have school-level authorities for making educational decisions. Data that we present below examines the kinds of authorities with responsibilities for making educational decisions and the kinds of decisions for which each has responsibilities. These data make clear considerable differences in how nations organize educational decision making. The locus of authority and how authority is exercised in decision making are related but show considerable variety. Thus, authority and educational decision making are a second kind of societal institutionalization examined here.

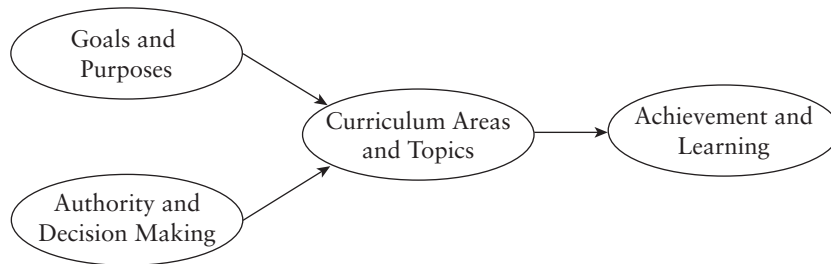
Curricular areas and topics are affected by goals and purposes and by authority and decision making as they are made real in the social institutions of various countries. For example, in mathematics in some countries proportionality is a separate content area treated directly in documents and textbooks and by teachers. In other countries, this content exists only as a part of the topic of algebra rather than as a separate topic. Thus curricular area or content topic is a third social institution of education that

will be examined. Some reserve the term *topic* for a segment of subject matter content, while *curricular area* is used for how a topic is represented in specific curricula. We will not maintain that fine a distinction here. We will use *curricular areas* and *topics* to refer to the content and expected performances, for contents of teaching units whether stated in official curriculum documents or in textbooks or as recognized by teachers. In this sense, a topic may not correspond from country to country because they are conceptualized, sequenced, and delivered differently. However, the commonalities are most often sufficient to recognize a core similarity for these curriculum areas and topics. When necessary, essential dissimilarities will also be emphasized.

Fourth and finally, achievement (or learning as change in achievement) is a social institution that is defined in varying ways in various societies. It is a reflection of what is learned and what capabilities are developed through the educative process. It is measured in different ways. In some countries, teachers' judgments without the use of formal assessment techniques are used to evaluate student attainments, especially prior to formal, national tests at the end of schooling. In other countries (for example, the United States), formal evaluation techniques such as teacher-made and standardized tests are used intensively as part of the common educational experience. For the sake of using TIMSS data, "achievement" is defined as a common core of what is measurable through tests related to curricula. These tests demand the display of learned capabilities under timed conditions and may or may not be perceived by those taking the tests as the more familiar, higher-stakes tests that are institutionalized in their country.

There is variation in curriculum among countries, regardless of the indicator used to reflect curriculum. There is variation in achievement and learning (gains in achievement) regardless of how this is limited by the way in which achievement is measured. However, our hypothesis is that national culture has an impact on curriculum. We believe it also has an impact on learning. Apart from how culture has an impact on curriculum and learning separately, culture also has an impact on the relationship between the two. This remains a hypothesis but we believe that it is true because we consistently find differences between countries. The way in which curriculum is related to learning varies among national cultures. For example, how directly textbooks reflect the content standards or curricular intentions and how directly they are related to achievement varies among countries. More detail on this and other examples are reported in various chapters throughout this book.

Figure 1.1. Four Social Institutions of Education Model.



The interactions of these institutions are complex, as are their conceptualizations. For the purposes of investigation here, we will assume a simple model that links these institutionalized components of education (see Figure 1.1). We believe it is reasonable to assume that the way a society institutionalizes goals and purposes relates to the way it institutionalizes authority and decision making. We believe further that these two typically interact and that together they help to shape curriculum areas and topics as they are institutionalized in national and subnational curricula of various forms. Further, we believe that curriculum areas and topics affect how achievement is realized. For simplicity, we assume that goals and decision making work through curriculum to affect achievement. Recall that this is an attempt to indicate key aspects of culture's impact on education indirectly through educational social institutions. Direct, diffuse culture impact is presumed informally to maintain its continuous impact on students at the same time.

The question as it is posed in general form here is not, "Does culture affect learning?" That is assumed. Instead we ask, "Specifically, how does culture, through educational social institutions shaped by culture, affect educational practice and outcomes?" We believe this to vary among countries and education systems. In this book we set ourselves the task of exploring this variation and answering the specific question as part of an answer to the question of how curriculum affects learning. More detailed models flowing from Figure 1.1 come in later chapters.

NOTES

1. The popular interpretation of the Coleman report (Coleman et al., 1966) was that schools don't make a difference, a perception that continues to

persist despite challenges to this interpretation (Hanushek, 1997) and other analyses that contradict it (Wenglinsky, 1997). A similar conclusion regarding schools appears to be supported by the argument set forth in *The Bell Curve* by Herrnstein and Murray (1994). A recent review argues that international studies provide critical evidence of the importance of schools in students' learning and cites the recantation Coleman made of his earlier work upon a re-analysis of international data (Suter, forthcoming).

2. Statistics Canada, Canada's national statistical agency, consulted with researchers in each country to draft national sampling plans according to the published TIMSS documents (Wolfe & Wiley, 1992; Foy & Schleicher, 1994). In addition, they reviewed and approved each country's sampling plan and all stages of the data sampling. See Foy, Rust, and Schleicher (1996) for a detailed account.
3. In the design and implementation of data gathering on comparative achievement, consensus political methods and multiple viewpoints affected the actual form, choice, and placement of items.
4. See relevant chapters in the three technical volumes edited by Michael O. Martin and Dana L. Kelly (1996, 1997, and 1998). These are available online at <http://times.bc.edu>.