

Levels of Comparison in Educational Studies: Different Insights from Different Literatures and the Value of Multilevel Analyses

MARK BRAY

University of Hong Kong

R. MURRAY THOMAS

University of California, Santa Barbara

Research in the field of comparative education has traditionally focused on studies across world regions and countries. In this article, Mark Bray and R. Murray Thomas argue that this approach often yields incomplete and unbalanced perspectives on educational studies because it fails to consider salient differences among states, districts, schools, classrooms, and individuals. The authors illustrate the need for a broader conceptualization of comparative education and propose multilevel analysis as a more comprehensive model of research that can integrate insights gleaned from single-level approaches.

Comparison is the basis of almost all forms of inquiry. However, the field of educational studies known as comparative education has historically been defined in a somewhat limited way. As active contributors to this field for some time, participating in its major journals, conferences, and professional bodies, we have felt disquiet about some of its basic premises, especially concerning the dominant units of comparison in the mainstream literature. We have noted that the chief focus in comparative education literature has been on countries and world regions, and that this has tended to lead to unbalanced and incomplete perspectives. On the other side of the coin, we have noted that much research in other fields of education has been undesirably localized in focus. In many of these studies, unbalanced and incomplete perspectives have resulted from the lack of an international dimension.

Harvard Educational Review Vol. 65 No. 3 Fall 1995
Copyright © by President and Fellows of Harvard College
0017-8055/95/0800-472 \$1.25/0

To analyze and investigate this situation more fully, we constructed a model for classifying comparative studies by level and type. We then examined a broad body of both quantitative and qualitative literature in education to select examples of studies focusing on local, regional, national, and international levels in a variety of contexts. Next, we identified the implications of those foci for interpretation and understanding of the subjects.

After considering studies that exclusively or mainly focus on single levels, we turned to multilevel analysis. Such analysis has become more common in recent years, partly because of increased awareness of its importance and partly because of the availability of new tools for research. However, a great deal of multilevel analysis remains at the individual, classroom, school, and perhaps district levels. We argue that the addition of state/provincial, national, and world-region comparisons would enhance understanding in some of these studies.

Finally, we observed that the research at particular levels is more common in some fields of educational studies than in others. For example, much of the work on effective schools has focused on institutions and what goes on inside them, and has made inadequate use of insights obtainable from cross-national comparison. The corollary is that most people working in the field commonly labelled comparative education tend to have a good understanding of macro-level phenomena but are much less comfortable with the tools and perspectives of researchers who work at the micro-level. We believe that both sides could learn from each other. We also highlight the value of multilevel analyses within the domain of educational studies.

A Framework for Comparative and Multilevel Analyses

For the purposes of this article, comparative education refers to all studies that inspect similarities and/or differences between two or more phenomena relating to the transmission of knowledge, skills, or attitudes from one person or group to another. This definition is deliberately broad, and embraces many foci and types of inquiry that would not normally be found in the pages of mainstream comparative education journals and textbooks. This is because the field of comparative education has chiefly developed from a concern with cross-national comparison, which remains its dominant focus.¹

In Figure 1 we present a three-dimensional way of classifying comparative studies. The first dimension, which will be the chief focus of this article, is geographic/locational. We identify seven levels: world regions/continents, countries, states/provinces, districts, schools, classrooms, and, finally, individuals.² A

¹ The content of the field is evident in its major journals and textbooks. The three major journals explicitly devoted to the field of comparative education are *Comparative Education Review*, *Comparative Education*, and *Compare*. Examples of major comparative education textbooks are Altbach, Arnove, and Kelly (1982); Arnove, Altbach, and Kelly (1992); Bereday (1964); Halls (1990); Ignas and Corsini (1981); Noah and Eckstein (1969); and Thomas (1990).

² It would not be difficult to identify further intermediate levels. For example, schools within districts may be grouped into clusters, teachers within schools commonly work in departments, and pupils within classrooms may work together as groups. However, the set of seven levels is adequate to make the main points of this article.

second dimension represents nonlocation demographic groupings, including ethnicity, religion, age, and gender, as well as an entire population. The third dimension embraces aspects of education and of society, such as curriculum, teaching methods, finance, management structures, political change, and labor markets. Every comparative study involves all three dimensions, and thus can be located in one or more of the cells in the diagram. As an example of how a study can be located in Figure 1, the shaded cell represents a research project comparing curriculum plans for all varieties of educational programs (entire population) in two or more provinces.³

Examples of Comparisons within Different Levels

In an article of this length, it is impossible to explore the full diversity of possible studies in the framework represented by Figure 1. To focus our discussion, the following examples concentrate on the seven levels along the geographic/location dimension. We present a few examples from recent literature for each level. The purpose of the examples is not only to illustrate how the framework may be applied, but also to expose the strengths and limitations of each level in terms of what it reveals about the entities being compared. For convenience, it is useful to describe the levels as if they were mutually exclusive, though in practice many studies embrace two or more tiers.

World Regions/Continents

A substantial literature focuses on the nature of educational provision in different regions of the world. Typical terms identifying regions are the Balkan States, the European Community, the Caribbean, and the South Pacific. Allied macro-level work takes the continent as the unit of analysis and focuses on such locations as Africa, South America, or Asia.

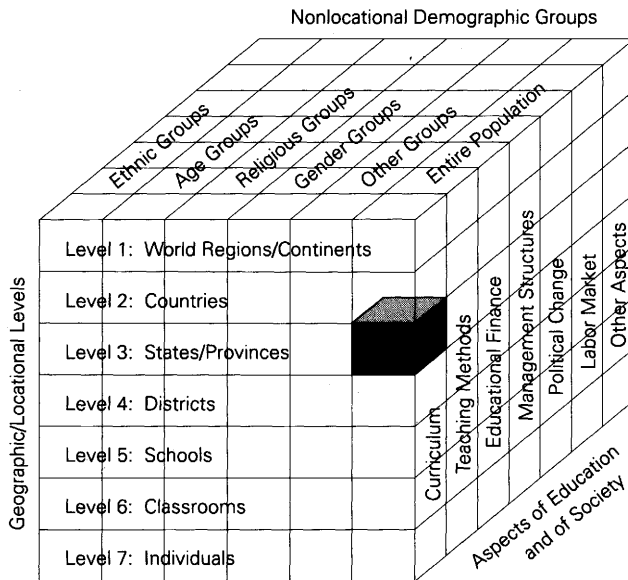
A key assumption underlying most regional comparisons is that certain shared characteristics differentiate one region from another in educationally important ways. The unifying characteristics of any particular region may include language, political organization, colonial history, economic system, national ambitions, and/or cultural origins. Three particular challenges face authors of cross-regional comparisons. They must convince readers that the characteristics cited as unifying a region are truly shared by the region's members; demonstrate that two or more regions are substantially similar or different in the nature of their unifying features; and show that such similarities and differences are educationally important.

To illustrate the potential benefits and drawbacks of analysis at the level of world regions, we present statistics from a study of official school curricula by Kamens and Benavot (1992) in Table 1. The authors collected information on curricula in a large sample of countries, and this particular table shows the

³ A further dimension of comparison would be across time. Thus, a phenomenon located in one particular cell in Figure 1 could be compared with the same phenomenon in the past and/or future. However, this dimension of comparison is beyond the focus of this article.

FIGURE 1

A Framework for Comparative Education Analyses



percentage of the primary curriculum devoted to mathematics and science in three historical periods. The authors, who are major figures in what is known as world systems analysis, concluded that there is strong global convergence in many curriculum components. Possible interpretation of the table is best presented in the authors' own words. The table, they suggested, shows that over time official attention to science and mathematics

grows in all regions (with the exception of Eastern Europe where trends are uneven). Across regions in the same time period, the differences are small, but there are a few interesting patterns. First, countries in Eastern Europe — and, to a lesser extent, in Latin America — place a greater emphasis on mathematics education than countries in other regions. In Eastern Europe this emphasis on mathematics apparently preceded their transition to socialism in the 1940s. Second, countries in Latin America devote a greater proportion of their curriculum to science than other regions, especially in the post-World War I period. Third, contrary to widely held impressions in both the popular and scholarly press, there is no evidence that Asian countries devote disproportionately more time to math and science education than other countries. And fourth, no region seriously devalues mathematics and science education. By and large, curricular variation by world region is rather limited in these two subject areas. (p. 119)

These observations are certainly instructive. From a methodological viewpoint, however, various factors might stress a need for caution in interpreting the data. First, it is interesting to note the construction of the world regions,

TABLE 1

Percentage of the Primary Curriculum Devoted to Mathematics and Science by World Region and Historical Period (Cases in Parentheses)^a

| | WORLD REGION | | | | | | TOTALS |
|--------------------|-----------------------|-----------------------------|--------------|-----------------------------|-------------------|--------------------------|--------------|
| | Sub-Saharan Africa | Middle East/ N.Africa | Asia | Latin America/ Carib. | Eastern Europe | The West ^b | |
| <i>Mathematics</i> | | | | | | | |
| 1920–1944 | 0.0 (1) | 14.2 (6) | 16.1 (5) | 16.5 (5) | 18.0 (5) | 15.6 (20) | 15.5 (42) |
| 1945–1969 | 15.1 (15) | 14.9 (12) | 15.1 (15) | 18.1 (14) | 21.9 (8) | 16.4 (17) | 16.5 (81) |
| 1970–1986 | 17.5 (15) | 16.6 (12) | 17.7 (15) | 19.2 (14) | 20.5 (8) | 18.5 (17) | 18.2 (81) |
| <i>Science</i> | | | | | | | |
| 1920–1944 | 0.0 (1) | 4.9 (6) | 4.7 (5) | 6.8 (5) | 9.4 (5) | 5.5 (20) | 5.8 (42) |
| 1945–1969 | 5.2 (15) | 5.7 (11) | 8.9 (14) | 9.2 (13) | 7.8 (8) | 6.3 (15) | 7.1 (76) |
| 1970–1986 | 7.1 (15) | 6.7 (11) | 7.9 (14) | 10.9 (13) | 7.7 (8) | 6.5 (15) | 7.8 (76) |

^a In order to increase the validity of longitudinal comparisons in this table, mean figures for each region in the 1945–1969 and 1970–1986 periods are calculated from a constant case base.

^b The 'West' region includes Western Europe, North America, Australia, and New Zealand.

Source: Kamens & Benavot (1992), p. 119.

each of which may appear appropriate on the surface, but most of which may be open to challenge. Thus, while it seems logical to separate Sub-Saharan Africa from North Africa, and to group the latter with the Middle East, the presentation of Sub-Saharan Africa as a single category means that former British, French, Portuguese, Belgian, and Dutch colonies are grouped together. Kamens and Benavot (1992) recognize that previous scholarship "suggests that the content of school curricula is related systematically to political, cultural and economic characteristics of each nation" (p. 118), including colonial history. The grouping of all countries of Sub-Saharan Africa into a single category, then, might obscure significant intra-continental diversity. Likewise, one might question the grouping of Latin America (which is mostly Spanish- and Portuguese-speaking) with the Caribbean (which includes English-, French-, and Dutch-speaking territories). Also, Asia includes countries as diverse as Mongolia, China, and Singapore; and if "the West" includes Western Europe, North America, Australia, and New Zealand, one might wonder if a similar case can be made for grouping former British colonies across the world (i.e., including Asia, Africa, and the Caribbean), or for grouping all Islamic countries (i.e., not just the Middle East and North Africa, but also such countries as Brunei Darussalam and Malaysia).

It is likely that Kamens and Benavot (1992) considered such possibilities, and the article from which the table is taken does include other tables that regroup the countries according to their level of economic development (i.e., more developed or less developed) and according to the period in which they achieved sovereign independence. These categories are also not without limitations, since the dividing line between more and less developed is arbitrary, and since transition to sovereignty does not always lead to great changes in dependency relationships. Nevertheless, comparison by world regions can be useful. Kamens and Benavot (1992) pointed out that

while more-developed nations place greater emphasis on mathematics in the primary curriculum than less-developed countries (the difference between them having narrowed in the recent period), this is not the case in the area of science. Surprisingly, in the latter two post-World War II periods, poor, agrarian Third World nations devote slightly more curricular time to science than more-developed nations. This negative relationship between level of economic development and curricular attention to science has also been shown in other longitudinal research designs. (p. 119)

This is not the place to explore in detail the findings of Kamens and Benavot and the extent to which their conclusions are justified by the data. However, this general discussion of their study indicates the limitations of analysis at the level of world regions.

Countries

Use of the country as the unit of analysis is quite common in the field of comparative education. Some textbooks (e.g., Fafunwa & Aisiku, 1982; Ignas & Corsini, 1981; Mazurek & Winzer, 1994) are completely organized around countries, presenting information in chapters simply labelled Afghanistan, Albania, Algeria, Australia, Austria, and so on. Many articles also take the country as the basic unit of analysis. As noted above, this is chiefly the result of emphasis within comparative education on cross-national comparison of systems. Such emphasis dates back to early seminal works in the field (e.g., Bereday, 1964; Hans, 1950; Kandel, 1933).

Table 2 reproduces data from a paper on rates of return to investment in education. Without entering the debate on the methodology for calculating and interpreting rates of return, which may itself be controversial (e.g., Klees, 1989; Leslie, 1990), some comments can again be made about comparative education methodology. While the table shows great variation both between countries and between rates of return to education at different levels, it seems to imply homogeneity within countries. This implication is misleading, for even small countries may have great diversity according to occupations, age groups, and regions.

A similar point may be made about Table 3, which presents information on literacy rates in different countries. The author noted that literacy rates in many of the countries shown were higher in urban than in rural areas (Tilak, 1994, p. 44). However, rural/urban and regional differences are not revealed by these national aggregates. Moreover, the fact that each country occupies the same

TABLE 2
Rates of Return to Education, by Country (%)

| Country | Year | Social | | | Private | | |
|---------------|------|---------|-----------|--------|---------|-----------|--------|
| | | Primary | Secondary | Higher | Primary | Secondary | Higher |
| Brazil | 1989 | 35.6 | 5.1 | 21.4 | 36.6 | 5.1 | 28.2 |
| Chile | 1989 | 8.1 | 11.1 | 14.0 | 9.7 | 12.9 | 20.7 |
| Ecuador | 1987 | 14.7 | 12.7 | 9.9 | 17.1 | 17.2 | 12.7 |
| Great Britain | 1978 | | 9.0 | 7.0 | | 11.0 | 23.0 |
| India | 1978 | 29.3 | 13.7 | 10.8 | 33.4 | 19.8 | 13.2 |
| Mexico | 1984 | 19.0 | 9.6 | 12.9 | 21.6 | 15.1 | 21.7 |
| Philippines | 1988 | 13.3 | 8.9 | 10.5 | 18.3 | 10.5 | 11.6 |
| United States | 1987 | | 10.0 | 12.0 | | | |
| Venezuela | 1989 | 23.4 | 10.2 | 6.2 | 36.3 | 14.6 | 11.0 |
| Zimbabwe | 1987 | 11.2 | 47.6 | 4.3 | 16.6 | 48.5 | 5.1 |

Source: Extracted from Psacharopoulos (1994), pp. 1340–1341.

amount of space (one line) in the table seems to imply equivalence among these countries. The table outlay glosses over the facts that national boundaries are entirely arbitrary, and that the forces of geography, history, and politics happen to have created units of greatly differing size and content.

Lest it be thought that the advantages and disadvantages of taking countries as the unit of analysis apply only to quantitative tables, an example from more qualitative research may also be presented. Many international agencies publish qualitative reports containing information on a country-by-country basis. One typical example is a 1984 UNESCO report on textbooks that presents information not only country-by-country, but also with a strong degree of uniformity. In this report, the subheadings were standardized, and the People's Republic of China was allocated 11 pages, while Fiji was allocated 12, India 16, Malaysia 10, Nepal 12, Pakistan 12, Papua New Guinea 11, Singapore 10, Thailand 11, and Western Samoa 12.⁴

On the positive side, these chapters certainly contain instructive information. Moreover, given the existence in many countries of national education systems that are strongly influenced by central government policies, it is legitimate to present at least some information in this way.

Again, however, the format seems to imply that the countries were equivalent units. Further, the format appears to overlook the fact that the country is not always the most appropriate unit for this type of information and analysis. India, for example, has a highly decentralized education system in which many state government policies are more important than national government ones (Bor-

⁴ Page lengths of other chapters varied considerably. For example, Indonesia was allocated 26 pages, Philippines 19, and Tonga 29. This variation was more a function of detail than a reflection of the complexity of circumstances in these countries or their relative size. It is striking that Tonga (population 95,000) was given more pages than Indonesia (population 179,000,000).

TABLE 3

Adult Literacy Rates in Selected Asian Countries, 1990 (%)

| | |
|-------------|----|
| Afghanistan | 29 |
| Bangladesh | 35 |
| China | 73 |
| India | 48 |
| Indonesia | 77 |
| Myanmar | 81 |
| Philippines | 90 |
| Thailand | 93 |
| South Korea | 97 |
| Japan | 99 |

Source: Tilak (1994), p. 194.

dia, 1988; Dyer, 1994). In other settings, by contrast, the international forces of dependency exerted through colonial traditions, examination systems, copyright, and multinational publishers are much stronger determinants of the nature of textbooks than are the policies of either national or sub-national governments (Altbach, 1992; Altbach & Kelly, 1988).

States/Provinces

Comparison of states or provinces is especially important in countries with high degrees of decentralization to the state/provincial level. In countries such as India, Australia, Canada, Germany, and the United States, state/provincial authorities may have a strong control over the structure of education, minimum qualifications for teachers, financial disbursements, curriculum, and other matters. It is thus difficult to talk meaningfully of *the* Australian system of education, for example. The same point applies to most other countries with federal structures.

Lawton's (1987) study of financial arrangements in Canada demonstrates the diversity within that country. Lawton observed dramatic variations in the methods by which elementary and secondary schools were financed in each of the ten provinces and two territories. For instance, whereas in Prince Edward Island all finance was provided by the provincial governments, in most other provinces, varying amounts were paid by local school boards. Lawton noted in passing that provincial governments had divergent policies with regard to Catholic and other church schools; and while in most provinces the principal medium of instruction was English, in Quebec the principal medium was French.

Of course, inter-provincial comparisons may also have different foci. Table 4 shows enrollment rates by province in Papua New Guinea. The figures were used to help the national government identify disadvantaged provinces in need of special assistance to achieve universal primary education. Once the enrollment rates had been determined, planners were in a better position to identify the reasons for inter-provincial disparities and to design tools to tackle the problem.

TABLE 4

Primary School Enrollment Rates, by Province, Papua New Guinea, 1983 (%)

| | | | |
|--------------------------|------|----------------------------|------|
| Southern Highlands | 46.8 | Oro | 65.3 |
| Eastern Highlands | 51.5 | West New Britain | 66.5 |
| Western Highlands & Enga | 52.6 | Gulf | 67.4 |
| Madang | 53.0 | North Solomons | 68.4 |
| Simbu | 53.4 | Western | 69.1 |
| Morobe | 53.5 | East New Britain | 75.2 |
| West Sepik | 54.8 | Central & National Capital | 76.7 |
| East Sepik | 59.5 | New Ireland | 79.2 |
| Milne Bay | 64.7 | Manus | 85.6 |

Source: Roakeina (1984), p. 33.

TABLE 5

Primary School Enrollment Rates by District, East New Britain Province, Papua New Guinea, 1983 (%)

| | | | |
|---------------|------|----------------|-------|
| West Pomio | 17.0 | Duke of York | 80.2 |
| Central Pomio | 52.1 | Lassul Baining | 85.7 |
| East Pomio | 63.3 | Melkoi | 91.6 |
| Bitapaka | 63.6 | Toma | 91.7 |
| Reimber | 69.6 | Livuan | 99.4 |
| East Baining | 72.3 | Raluana | 104.2 |
| Watom | 73.3 | Kombiu | 105.5 |
| Balanataman | 76.5 | Vunamami | 113.5 |
| Rabaul Town | 78.0 | Central | 117.5 |

Note: The figures show gross enrollment rates, i.e., the number of children in primary school divided by the number of children in the age group. The chief reason why some figures exceed 100% is that children older and/or younger than the official age group were enrolled in primary schools in those (and perhaps also in some other) districts.

Source: East New Britain Provincial Government, Rabaul.

However, like Tables 2 and 3, which took countries as the unit of analysis, Table 4 obscures major differences within provinces.

Districts

To demonstrate the extent to which the Papua New Guinean provincial data on primary school enrollment rates obscured lower-level disparities, Table 5 presents figures on districts within the province of East New Britain. Although the inter-provincial statistics in Table 4 had suggested that the province of East New Britain had a relatively high enrollment rate, the district-level data in Table 5 show considerable variation, with West Pomio having a notably low enrollment rate of 17.0 percent. These and other district-level disparities have been of major concern to the provincial authorities (Weeks & Waninara, 1988), and arguably should also concern the national government.

An alternative example of district-level analysis is a recent study of basic education in three counties in China (Lewin & Wang, 1994). The rationale for the study, as stated at the beginning of the report, was that

national statistics on basic education are an unsatisfactory guide to policy and decision making in China. Though the overall picture of impressive achievement appears largely justified we are aware that aggregated statistics at the national-level may be unreliable. (p. 1)

The report indicated that the unreliability of national statistics was partly due to the technical difficulties of data aggregation across widely varying circumstances, the uncertainties inherent in data collection, and the fact that some statistics were clearly undependable. The report added that

national statistics only provide one perspective on the extent and quality of the implementation of basic education policy. Judgements on the effectiveness of policy implementation require insights that can only be obtained from case study work at the local-level [sic]. It is here that the decisions are made which affect enrollment, drop-out, repetition, resource allocation, teaching quality, and achievement and it is here that an understanding of policy in action is assessable. (p. 2)

The researchers selected counties with contrasting conditions to highlight the need for disaggregated data. One was a prosperous county near Beijing, another was an economically disadvantaged county with a majority Han population, and the third was a particularly poor county with a large minority population. The researchers showed that different policies for educational administration and management were needed for each type of situation. This would not have been evident from more aggregated data.

Schools

The literatures that take schools as the unit of analysis are rather different from those that focus on countries, provinces, and districts. Analysis at these higher levels of governance may be concerned with the people who are *not* enrolled in schools as well as with those who are. Such a focus is not possible with literature that takes schools as the unit of analysis. Moreover, adoption of the school as the unit of analysis requires a focus on institutional culture, which is rather different from the cultural focus underlying larger units.

The book *Fifteen Thousand Hours* by Rutter, Maughan, Mortimore, and Ouston (1979) is one well-known study that took the school as the unit of analysis. The researchers wished to investigate, among other questions, whether it mattered *which* school the children attended, and, if so, what features of schools made a difference in terms of academic achievement and social development. They selected twelve schools in inner London, where they examined the characteristics of the pupils and teachers, the aims of the administrators, the facilities, and various other factors contributing to the ecologies of each institution. The researchers concluded that each school did have its own ethos, which was partly built on the expectations of administrators, teachers, and pupils (p. 182). The study thus identified a form of institutional culture that was of considerable

importance both to the nature of education and to the broader shaping of society, and that could not have been identified if higher or lower units had been used for analysis.

An example from a rather different setting but also in the United Kingdom is Forsyth and Carter's (1983) study of the cultures of small rural schools. Again, their concerns were to investigate the operation of the schools as institutions, though their work contrasted with *Fifteen Thousand Hours* by including a strong focus on the influence of the wider community on those institutions.

One feature of this level of research is that it can present personalized portraits. Research at district, provincial, and national levels can, of course, include the impact of specific policymakers at the apex of each hierarchy, but it cannot bring into focus the impact of individual differences among the "ordinary" actors in the same way. Another important factor is that schools are sufficiently numerous to permit meaningful random sampling, which would not normally be possible at the world-region, national, or provincial levels, though it could in some contexts be appropriate at the district level.

Classrooms

As with the other levels, a considerable literature uses the classroom as the unit of analysis. For example, sociologists have examined classrooms as social systems, noting the ways that the pupils interact with each other and with their teachers.

One of the best-known qualitative studies is by Jackson (1968), who focused on elementary school classrooms in the United States. In *Life in Classrooms*, Jackson noted differences between classrooms, though at the outset he put more stress on similarities. For example:

In their efforts to make their classrooms more homelike, elementary school teachers often spend considerable time fussing with the room's decorations. Bulletin boards are changed, new pictures are hung, and the seating arrangement is altered from circles to rows and back again. But these are surface adjustments at best. . . . School bulletin boards may be changed but they are never discarded, the seats may be rearranged but thirty of them are there to stay, the teacher's desk may have a new plant on it but there it sits, as ubiquitous as the roll-down maps, the olive drab wastebasket, and the pencil sharpener on the window ledge. (pp. 6-7)

Other research that compares classrooms has focused on such factors as size, composition of student population, communication patterns, and reward structures. Boocock's (1980) textbook on the sociology of education is typical of many in that it contains a separate chapter to contrast the classroom as a social system with the school and the wider society. The point in Boocock's book, as in this article, is that similarities and contrasts across different levels of analysis generate different insights. Similarities stem primarily from the fact that each level of society contains people who relate to each other in predictable ways according to the parameters of that society. Contrasts, on the other hand, chiefly arise from the compositions of the societies at different levels and from the natures of the interactions that take place.

Anderson and Burns (1989, p. 41) make a related point, but contrast the classroom with the teacher or pupil as units of analysis. Their concern is to highlight the predictability of the activity structure of classrooms while at the same time stressing the multidimensionality and complexity of interaction patterns. Anderson and Burns indicate that teachers and students are often unaware of the dynamics that are linked to the classroom *as* a classroom. Research that takes the classroom as the unit of analysis thus permits the actors to understand better the forces that operate within the four walls, and the reasons for the behavior not only of the other people, but also of themselves.

Individuals

Research may also focus on individuals: principals, teachers, parents, pupils, and others. Such studies may have many disciplinary orientations, but are more likely than analyses at other levels to emphasize psychology.

One example of research at this level is the chapter in Biggs and Moore (1993) on students' approaches to learning. The chapter contrasts Student A, who is motivated to learn, has a deep approach to learning, and gets fully involved with the topic, with Student B, who wants to have fun, has a surface approach to learning, and decides to do the minimal amount, and Student C who wants top marks, has an achieving approach to learning, and uses time effectively (p. 306). Among the authors' goals were to show that:

- different students may perceive schooling in different ways;
- students can be made aware of how they learn and solve problems;
- aspects of the more effective approaches can be taught as learning/study skills; and
- for various reasons, teaching of study skills is not always effective.

An alternative example is the work of Toomey, Mahon, and Thalathoti (1993), who investigated the ways in which technological innovations changed the ways that teachers organized their lessons. The authors identified two basic types of teachers: 1) "interactionist" ones who usually integrated subject matter, gave pupils active roles, and encouraged cooperative learning, and 2) "interventionist" ones who usually separated subject matter, expected pupils to be passive, and emphasized individual work. The personalized reporting that is possible at this level may be illustrated by the views of two specific teachers presented in the research. Anthea, who is described as an interactionist teacher, is reported to have felt that

the real value of this exercise is that children work together, develop their own questions, do their own research, decide how to report their research. . . . They feel like they own the work they do . . . and more importantly they are deciding how to use the computer and making it work on their own terms. (p. 307)

In contrast, Lee is described as an interventionist teacher. She feels that her way of teaching is "first to come to grips with what it is I have to teach them, and to then give it to them . . . including things to do with computing" (p. 307).

Not all research that takes individuals as the unit of analysis can have this personalized flavor; in contrast are large-scale surveys of teachers, pupils, or other individuals. For example, the Hong Kong government's annual survey of teachers presents statistics on the gender, qualifications, ages, and other characteristics of teachers in a highly aggregated way (Government of Hong Kong, 1994). Of course, any interpretation of the statistics in this survey can only be undertaken in light of information about other components of the education system and society, but the survey is a valuable source of information for planners, administrators, trainers, and others.

Multilevel Analysis

The above examples primarily (though not exclusively) focus on single-level analysis. In contrast, various studies use a multilevel design in order to achieve more complete and balanced understandings. While many such studies suffer flaws of various kinds, the fact that they consider their subjects from several different angles facilitates more comprehensive and possibly more accurate presentation of the phenomena they address.

The importance of multilevel analysis for certain types of inquiry has been recognized increasingly during the last two decades (e.g., Burstein, 1980, 1988; Cronbach, 1976; Goldstein, 1987; Raudenbush & Willms, 1991). As summarized by Keeves and Sellin (1988),

Formerly the issue associated with the appropriate level of analysis was considered to be influenced largely by the nature of the research questions to which answers were sought, for example, whether the problem was concerned with individual students or with classroom groups, as well as by the level at which sampling had taken place and at which generalization to other situations was sought. More recently it has become apparent that a multilevel analysis strategy is required if appropriate answers are to be obtained. (p. 690)

However, the dominant form of research under the specific label of multilevel analysis has been principally confined to the individual, classroom, and school levels. Such studies have generally omitted careful consideration of the state/province, country, and world-region levels, with the result that interpretations have still been arguably unbalanced and incomplete, albeit more informative than before.

One area in which multilevel analysis has made a major contribution is school effectiveness. For example, in 1986, Raudenbush and Bryk presented a hierarchical model for studying school effectiveness in the United States. The authors began their paper with examples of other research on the effects of different interventions on student outcomes that have generated false conclusions because of failure to examine phenomena at different levels. Raudenbush and Bryk (1986, p. 1) suggested that most of these researchers clung to single-level models not so much out of conviction but because of the absence of viable alternatives. In this particular sub-field, increasing awareness of the mismatch between social

processes and the traditional statistical models used to study them has spurred a search for multilevel analytical strategies. Moreover, advances in statistical theory and increasingly powerful computers have provided the necessary tools. The remainder of Raudenbush and Bryk's paper focuses on one such tool, and shows how it can be used to advance understanding through disentangling school, classroom, and individual-level variables and contrasting them with each other in a more informative way.

In a similar vein, Kreft (1993) used multilevel analysis to assess determinants of school effectiveness in a sample of Dutch secondary schools. Addressing the same sorts of literature as Raudenbush and Bryk (1986), Kreft referred to the rise of so-called second-generation research in this sub-field, which, instead of making students the primary unit of analysis, recognizes that "students are nested in classrooms; classrooms are nested in schools; and schools are nested in school systems, in regions, in counties, or in countries" (Kreft, 1993, p. 106). One noticeable difference between the first and second generations of research, Kreft indicates, is a reduction of emphasis on externally determined factors such as teachers' salaries, per-pupil expenditures, the turnover of teachers, and physical facilities, and increased emphasis on the internal organization of schools.

Despite this shift of emphasis, however, most multilevel analysis in the field of school effectiveness has remained at the levels of school and below. The review of models by Bosker and Scheerens (1994) does mention the context within which schools operate, but gives it very little attention. Likewise, although the book edited by Raudenbush and Willms (1991) includes studies from Israel, the United States, Scotland, Thailand, and the Netherlands, and is subtitled "International Studies of Schooling from a Multilevel Perspective," it makes few direct comparisons across national boundaries, and fails to draw out the implications of national-level forces. Rather, the book is a collection of local-level studies from different countries, which are placed side by side in the same volume. Similarly, while Kreft (1993) included countries in her list of levels in her hierarchical nesting model, her paper was primarily concerned with schools, classrooms, and individuals, and placed little emphasis on system-wide factors.

The chief explanation for this gap seems again to lie in the different skills of individual researchers and in the paradigms of particular types of research. The model presented by Raudenbush and Bryk (1986) is highly quantitative and requires a computer to perform various kinds of multiple regression. Because of the smaller numbers that could be put in samples, data of the required type for such analysis are less likely to be available at the national and world-region levels. This is one reason why much cross-national comparative research is qualitative rather than quantitative. Such researchers, even when they address issues of school effectiveness, tend to be more concerned with the influences of politics and macro-culture on the shape of school systems, including such forces as the dependency of Third World countries on the products, technology, and intellectual fashions of industrialized nations (e.g., Levin & Lockheed, 1993). Few individual researchers are equally comfortable with the literature and concepts of these two very different emphases.

Given this fact, one way to overcome the limitations of individual researchers is to assemble teams. An example that demonstrates what can be achieved in this way is the investigation of reading literacy conducted under the auspices of the International Association for the Evaluation of Educational Achievement (IEA). The study focused on the reading skills of 210,000 pupils aged nine and fourteen, taught by 10,000 teachers in the mainstream schools of thirty-two countries (Elley, 1992, 1994; Lundberg & Linnakyla, 1993; Postlethwaite & Ross, 1992). Findings were analyzed at the country, district, school, class, and individual levels. On the nonlocational demographic axis, analyses focused on pupils, including their ages and genders, and also on teachers and parents. The types of reading on which pupils were tested included a) narrative passages ranging from short fables to lengthy stories; b) other forms of writing such as family letters, simple descriptions of animals, and elaborate descriptions of lasers and the dangers of smoking; and c) graphs, charts, maps, lists, and sets of instructions. Additional aspects of the educational enterprise inspected included hours of reading instruction, teachers' instructional methods, school and class sizes, pupil/teacher ratios, number of textbooks per class, and number of textbooks per school library. Among factors studied in the children's out-of-school environments were number of books in the public libraries and bookstores, and home conditions such as socioeconomic status, available reading materials, TV viewing, voluntary reading activities, and parental cooperation in fostering children's reading skills. Representation of the reading-literacy study in terms of Figure 1 would require the shading of many cells within the cube.

Had this study been conducted within a single country, it would of course have been possible to investigate all sorts of relationships between these various factors within that country. Yet, the fact that it was a thirty-two-country study meant that insights were multiplied much more than thirty-two times. Researchers in education are handicapped by the fact that for reasons of political constraints and human rights, many variables in specific countries have to be taken as fixed. However, when a wide range of countries is investigated, the researchers are in effect presented with a natural laboratory containing a wider range of variables. For example, average class size for fourteen-year-old pupils varied from forty-eight in the Philippines to just eighteen in Switzerland, and the percentage of pupils who normally spoke a different language from the language of instruction ranged from 82 percent in Singapore to zero in Finland. Further, the proportion of teachers who were female ranged from 97 percent in Slovenia to just 21 percent in Switzerland. The sample also included relatively poor countries, such as Nigeria, and relatively rich ones, such as Iceland. By looking at the situations in other countries, researchers in their home countries were stimulated to ask questions they would never have asked had they been working in isolation.

This IEA study was not without problems, however. One of the most obvious difficulties arose from the demands of collecting and analyzing so much data. The project required the cooperation of many people, great amounts of time, powerful computers, and diverse kinds of expertise. What researchers gain in

producing such a sophisticated, multilevel, and multi-dimensional understanding of a phenomenon must be paid for; and in most countries, it has become more difficult than before to secure resources for such research. From one perspective, this is regrettable, but it does not mean that meaningful multilevel research cannot be carried out. The question of multilevel research design is more a matter of paradigm and philosophy than scale.

Conclusions: Different Levels, Different Insights

Our starting point in this article was the field of comparative education, which we argue should be recognized as legitimately including much more than world-region and cross-national studies. We suggest that until now, the field as traditionally defined has been too exclusive in focus. At the same time, too many studies in other fields lack an international dimension. They may also suffer from a lack of multilevel analysis, even within a national framework.

The strength of studies on the upper levels of the hierarchy (i.e., state/province, country, and world-region) is that they can offer general frameworks within which to place the more specific details of particular settings. Such frameworks can provide an initial basis for understanding and interpretation, and reduce the danger of overwhelming researchers and their audiences with masses of particularistic detail. Research focusing on the highest levels also helps identify broad economic conditions, political structures, cultural traditions, and forms of educational organization and administration that influence how much of what type of education is provided for different sectors of society.

The weakness of such work, however, is that the broad generalizations obscure the features that distinguish one region, school, or pupil from another. Macro-level studies cannot recognize individual differences, or account for the importance of those differences in educational events. Moreover, while it is useful in some contexts to group countries into categories labelled more developed or less developed, or by continent or world-region, in other contexts this grouping may be problematic. We have suggested in this article that the work of Kamens and Benavot (1992), for example, is important in clarifying the nature of curriculum change, but that it is misleading insofar as it portrays countries or education systems as homogeneous entities and thereby overlooks many of the local factors that may provide better explanations of curriculum change in particular settings.

A corresponding point is that much literature on lower levels (e.g., individual, classroom, and school) makes inadequate use of higher level studies. For example, the value of the research on students' approaches to learning reported in Biggs and Moore (1993) can only be maximized if it is seen in the context of broader factors that shape both the orientations of teachers and students and the nature of classrooms and schools. In countries with strong state/provincial variation, comparisons at that level can stimulate researchers whose primary focus is at a lower level to ask questions about features of their systems that they would otherwise have taken for granted. The same is true of cross-national com-

parisons, where differences are likely to be even more marked. Broader foci may also sharpen awareness of underlying forces. In this article, we have given the example of a country-by-country study of textbooks that made insufficient use of research on transnational influences and dependencies.

These observations lead to the conclusion that much, and perhaps most, research requires multilevel comparative analysis in order to achieve a full and balanced understanding of its subjects. Excellent investigations do exist that build on multilevel research to analyze subjects from different angles, but the number of such studies is much more limited than might be expected or desired.

The question then becomes why the number of such studies is so limited. Part of the answer lies in the fact that much research is conducted by individuals or, at most, by small teams, and is thus constrained by the perceptions, background, and training of those people. A second element lies in the fact that many fields within the macro-framework of educational studies are rather introspective, with individual researchers relating to each other within their fields but with little cross-fertilization from other fields. One example given in this article is of work in the field of school effectiveness, which, although multilevel, has been mainly focused on schools, classrooms, and individuals. Much of this work has been highly quantitative, and has not taken adequate account of the more qualitative contributions from the field of cross-national comparative education. Likewise, specialists in cross-national comparative education have tended to treat national systems of education as if they were largely homogeneous, and have made inadequate use of the comparisons conducted by specialists in other fields that expose considerable diversity among states/provinces, districts, schools, classrooms, and individuals.

These comments are not, of course, to deny that much research focusing on single levels is also valuable. However, researchers presenting such work would assist their readers if they called more attention to the limitations of their research. In their eagerness to inform the public about their findings, such researchers are prone to overlook alternative perspectives that have not been addressed and that might link to rather different interpretations. In this regard, the framework for multilevel analysis presented in this article can help identify the perspectives from which educational phenomena have and have not been investigated, and suggest a more balanced and comprehensive approach for future research.

References

- Altbach, P. (Ed.). (1992). *Publishing and development in the third world*. Oxford, Eng.: Hans Zell.
- Altbach, P., Arnove, R., & Kelly, G. (Eds.). (1982). *Comparative education*. London: Macmillan.
- Altbach, P., & Kelly, G. (Eds.). (1988). *Textbooks in the third world: Policy, content and context*. New York: Garland.
- Anderson, L., & Burns, R. (1989). *Research in classrooms: The study of teachers, teaching and instruction*. Oxford, Eng.: Pergamon Press.

- Arnové, R., Altbach, P., & Kelly, G. (Eds.). (1992). *Emergent issues in education: Comparative perspectives*. Albany: State University of New York Press.
- Bereday, G. (1964). *Comparative method in education*. New York: Holt, Rinehart & Winston.
- Biggs, J., & Moore, P. (1993). *The process of learning* (3rd ed.). Sydney: Prentice Hall.
- Boocock, S. (1980). *Sociology of education: An introduction*. Lanham, MD: University Press of America.
- Bordia, A. (1988). India. In T. Postlethwaite (Ed.), *The encyclopedia of comparative education and national systems of education* (pp. 350–358). Oxford, Eng.: Pergamon Press.
- Bosker, R., & Scheerens, J. (1994). Alternative models of school effectiveness put to the test. *International Journal of Educational Research*, 21, 159–180.
- Burstein, L. (1980). Analysis of multilevel data in educational research and education. In D. Berliner (Ed.), *Review of research in education* (vol. 8, pp. 158–233). Washington, DC: American Educational Research Association.
- Burstein, L. (1988). Units of analysis. In J. Keeves (Ed.), *Educational research, methodology and measurement: An international handbook* (pp. 775–781). Oxford, Eng.: Pergamon Press.
- Cronbach, L. (1976). *Research on classrooms and schools: Formulation of questions, design and analysis* (Occasional paper). Stanford, CA: Stanford Educational Consortium.
- Dyer, C. (1994). Education and the state: Policy implementation in India's federal polity. *International Journal of Educational Development*, 14, 241–253.
- Elley, W. B. (1992). *How in the world do children read?* The Hague: IEA.
- Elley, W. (Ed.). (1994). *The IEA study of reading literacy: Achievement and instruction in thirty-two school systems*. Oxford, Eng.: Pergamon Press.
- Fafunwa, A., & Aisiku, J. (Eds.). (1982). *Education in Africa: A comparative survey*. London: George Allen & Unwin.
- Forsyth, D., & Carter, I. (1983). The local view: A case study approach. In D. Forsyth, I. Carter, G. Mackay, J. Nisbet, P. Sadler, J. Sewel, D. Shanks, & J. Walsh (Eds.), *The rural community and the small school* (pp. 137–170). Aberdeen: Aberdeen University Press.
- Goldstein, H. (1987). *Multilevel models in educational and social research*. New York: Oxford University Press.
- Government of Hong Kong. (1994). *Teacher survey 1993*. Hong Kong: Education Department, Statistics Section.
- Halls, W. D. (Ed.). (1990). *Comparative education: Contemporary issues and trends*. Paris: UNESCO.
- Hans, N. (1950). *Comparative education: A study of educational factors and traditions* (2nd ed.). London: Routledge & Kegan Paul.
- Ignas, E., & Corsini, R. (Eds.). (1981). *Comparative education systems*. Itasca, IL: F. E. Peacock.
- Jackson, P. (1968). *Life in classrooms*. New York: Holt, Rinehart & Winston.
- Kamens, D., & Benavot, A. (1992). A comparative and historical analysis of mathematics and science curricula, 1800–1986. In J. Meyer, D. Kamens, & A. Benavot (Eds.), *School knowledge for the masses: World models and national primary curricular categories in the twentieth century* (pp. 101–123). London: Falmer Press.
- Kandel, I. (1933). *Studies in comparative education*. London: George G. Harrap.
- Keeves, J., & Sellin, N. (1988). Multilevel analysis. In J. Keeves (Ed.), *Educational research, methodology and measurement: An international handbook* (pp. 689–700). Oxford, Eng.: Pergamon Press.
- Klees, S. (1989). The economics of education: A more than slightly jaundiced view of where we are now. In F. Caillods (Ed.), *The prospects for educational planning* (pp. 244–291). Paris: International Institute for Educational Planning.
- Kreft, I. (1993). Using multilevel analysis to assess school effectiveness: A study of Dutch secondary schools. *Sociology of Education*, 66, 104–129.
- Lawton, S. (1987). *The price of quality: The public finance of elementary and secondary education in Canada*. Toronto: Canadian Education Association.
- Leslie, L. (1990). Rates of return as informer of public policy: With special reference to the World Bank and third world countries. *Higher Education*, 20, 271–286.

- Levin, H., & Lockheed, M. (Eds.). (1993). *Effective schools in developing countries*. London: Falmer Press.
- Lewin, K. M., & Wang, Y. J. (1994). *Implementing basic education in China: Progress and prospects in rich, poor and national minority areas* (Research Report No. 101). Paris: International Institute for Educational Planning.
- Lundberg, I., & Linnakyla, P. (1993). *Teaching reading around the world*. The Hague: IEA.
- Mazurek, K., & Winzer, M. (Eds.). (1994). *Comparative special education*. Washington, DC: Gallaudet University Press.
- Noah, H., & Eckstein, M. (1969). *Towards a science of comparative education*. London: Macmillan.
- Postlethwaite, T., & Ross, K. (1992). *Effective schools in reading: Implications for planners*. The Hague: IEA.
- Psacharopoulos, G. (1994). Returns to education: a global update. *World Development*, 22, 1325-1343.
- Raudenbush, S., & Bryk, A. (1986). A hierarchical model for studying school effects. *Sociology of Education*, 59(1), 1-17.
- Raudenbush, S., & Willms, J. D. (1991). *Schools, classrooms and pupils: International studies of schooling from a multilevel perspective*. San Diego: Academic Press.
- Roakeina, G. (1984). *A statistical review of the education sector*. Waigani, Papua New Guinea: Ministry of Education.
- Rutter, M., Maughan, B., Mortimore, P., & Ouston, J. (1979). *Fifteen thousand hours: Secondary schools and their effects on children*. London: Open Books.
- Thomas, R. M. (Ed.). (1990). *International comparative education: Practices, issues, and prospects*. Oxford, Eng.: Pergamon Press.
- Tilak, J. (1994). *Education for development in Asia*. New Delhi: Sage.
- Toomey, R., Mahon, L., & Thalathoti, V. (1993). The Linkwest Project: Comparative case studies in computer enhanced teaching and learning. *Journal of Technology and Teacher Education*, 1, 299-313.
- UNESCO. (1984). *Textbooks and reading materials*. Bangkok: UNESCO Regional Office for Education in Asia and the Pacific.
- Weeks, S., & Waninara, J. (1988). *Review of the education system in East New Britain*. Waigani, Papua New Guinea: National Research Institute, Educational Research Division.