7

The Evolution of Comparative and International Education Statistics

Stephen P. Heyneman and Chad R. Lykins

INTRODUCTION

An argument ensued between M. Dottrens and Jean Piaget at a board meeting of the Institute of Statistics in Geneva in 1933.[1] Dottrens proposed an international survey to record what countries were doing in education. Piaget was against it. He said: "L'expérience nous a montre qu'il est extrêmement difficile d'établir des tableaux statistiques comparables" (Smyth, 1996, p. 4). Piaget had a point. At that time there was no common definition on what education meant, how schooling might differ from ad hoc learning, or how to distinguish educational levels. The meanings of vocational and general education varied between and within nations. There were 115 different ways to define literacy and 133 different ways to classify educational attainment by age group (Smyth, 2005, p. 13). Dottrens, however, apparently won the argument on grounds that, in spite of the procedural complexities and the danger of receiving misleading results, the demand to know what countries are doing in education was simply irresistible.

These same arguments have re-occurred with regularity in the last 74 years. Though those who agree with Dottrens' claim have lost many battles in the interim, they have all but won the war. The record of advancement in geographical coverage and qualitative depth in comparative education statistics has been unidirectional. From counting schools in 1933 to videotaping teaching styles and capturing unit expenditures, the story of educational measurement and the resulting debates over its unprecedented findings is one true sign that there has been progress in education research. This extraordinary growth, both in the quantity and the quality of educational data, has brought fresh—and sometimes contentious—insights into perennial questions concerning the financing of education. The goal of this chapter is to explore the history of this growth and what it means for the study of some key questions in education finance.

EDUCATION COMPARISONS IN OECD COUNTIES

The Origin of Descriptive Education Statistics

The first attempt to compare educational expenditures internationally was made in 1937 by the International Bureau of Education (IBE). Their report presented seven tables illustrating national budgets, number of primary and secondary schools, and teachers and students in 58 countries. In

May of 1946, the United States submitted to the UNESCO Preparatory Commission a proposal to establish an international statistical service which should "assist in the co-ordination, standardization and improvement of national education statistics, the technology of its standardization, and should provide advice to member countries and international organizations on general questions relating to collection, interpretation and dissemination of education statistics" (Smyth, 1996, p. 7). The proposal was adopted. Consequently, a meeting of experts was called in November of 1951, to work on details of the world's first education statistical service.

The barriers to its operations were substantial. Prior education statistics consisted of only country-by-country "illustrations." There was insufficient agreement on terms and procedures to make comparisons. Recognizing this lack of agreement, the committee developed the first International Standard Classification of Education (ISCED) document which defined a school and established standards for classifying primary, secondary, higher and vocational education in 1953. It also provided the first definitions of education finance. These included definitions for *receipts* (cash received or made available for schools, including appropriations, subventions, fees, cash value of property received as gifts), *expenditures, recurring expenditures*, *capital expenditures* and *debt service* (Smyth, 1996, p. 18). Under each category, moreover, there were lengthy descriptions of each concept and details concerning proper tabulation. These initial iterations resulted in the compendious *World Surveys of Education* published in 1955, 1958, 1961, 1966, and 1971.

Despite significant improvements in education statistics between 1933 and 1971, there were still no indicators of education achievement or outcomes. Nor were there reliable indicators of teaching technologies, student attitudes, labor market activities, or unit expenditures. This lack became the focus of international attention and debate beginning in the 1970s and continuing to the present.

The Origin of Cross National Tests of Academic Achievement

The effort to compare academic achievement across countries was among the most controversial aspects of international education statistics. At first, the use of surveys for cross national education research was just an experiment, born from a chance visit by Torsten Husen (from the University of Stockholm) to the University of Chicago in the mid-1950s. There Husen met Professors Benjamin Bloom (Curriculum), C. Arnold Anderson (Sociology) and Mary Jean Bowman (Economics). They agreed that the world should be seen as a single educational laboratory. From this meeting emerged the idea—novel at the time—to sponsor an informal academic association promoting surveys of education *achievement*. The group was called the International Association for the Evaluation of Educational Achievement (IEA). For reasons of diplomacy, it was headquartered in Sweden. In 1959, the IEA published the first results from a pilot survey of a test consisting of 120 items covering reading comprehension, geography, science, mathematics, and non-verbal ability administered to non-representative samples of 9,918 13-year-old students in 12 countries (Postlewaite, 1975, p. 1). Since then there have been 33 cross-national studies of academic achievement, 29 of which have been sponsored by the IEA (Figure 7.1.).

Problems emerged from the beginning. It was a significant challenge for academics to manage an enterprise of such political and institutional complexity. There were difficulties in agreeing on common definitions, methodologies, sampling, and techniques for managing data. There were hurdles in obtaining sufficient resources from participating countries and donors with widely divergent economic circumstances. And there were severe pressures to meet strict schedules of implementation. These spawned numerous assessments of the benefits and drawbacks of the cross-national enterprise (Heyneman, 2004a, 2004b; Chromy, 2002; Floden, 2002; Linn, 2002;

Sponsor	Description	Countries	Year(s) Conducted
IEA	First International Mathematics Study (FIMS)	12 countries	1964
IEA	Six Subjects Study:		1970–1971
	Science	19 systems	
	Reading	15 countries	
	Literature	10 countries	
	French as a foreign language	8 countries	
	English as a foreign language	10 countries	
	Civic Education	10 countries	
IEA	First International Science Study (FISS) (part of Six Subjects Study)	19 countries	1970–1971
IEA	Second International Mathematics Study (SIMS)	10 countries	1982
IEA	Second International Science Study (SISS)	19 systems	1983–1984
ETS	First International Assessment of Educational Progress (IAEP-I, Mathematics and Science)	6 countries (12 systems)	1988
ETS	Second International Assessment of Educational Progress (IAEP-II, Mathematics and Science)	20 countries	1991
IEA	Reading Literacy (RL)	32 countries	1990–1991
IEA	Computers in Education	22 countries	1988–1989
		12 countries	1991-1992
Statistics International Adult Literacy Survey (IALS) Canada		7 countries	1994
IEA	Preprimary Project:		
	Phase I	11 countries	1989–1991
	Phase II	15 countries	1991–1993
	Phase III (longitudinal follow up of Phase II sample)	15 countries	1994–1996
IEA	Language Education Study	25 interested countries	1997
IEA	Third International Mathematics and Science Study (TIMSS):		
	Phase I	45 countries	1994–1995
	Phase II (TIMSS-R)	About 40	1997–1998
IEA	Civic Education Study	28 countries	1999
OECD	Program for International Student Assessment	32 countries	2000 (reading) 2003 (math) 2006 (science)

FIGURE 7.1 Selected International Comparative Studies in Education: Scope and Timing.
Source: Chromy (2002).

Rowen, 2002; Smith, 2002; Postlewaite, 1999; Goldstein, 1995; Horvitz, 1992; Medrich and Griffith, 1992; Olkin and Searls, 1985).

The surveys did, however, help stimulate debate over new questions and theories from a wide variety of disciplines including sociology, philosophy, child and cognitive psychology, economics, pedagogy, and education policy (Medrich and Griffith, 1992; Purves and Levine, 1975; Tomlinson and Tuijnman, 1994; Phillips, 1991; IEA, 1990). For instance, Robert Thorndike observed that the variables which differentiate students are not the same as those which differentiate countries. On the basis of international comparisons, he once argued that parental help may be a sign of student ineptitude (Thorndike, 1975, p. 102). Benjamin Bloom noted that students

who do poorly are those who are not interested in studying (Bloom, 1975, p. 80). Contrary to the theory that elite education systems would have higher achievement, William Platt found that achievement does not decline as universal schooling and rates of retention go up. He also found that achievement stems from certain cognitive domains rather than others (Platt, 1975, p. 40). Each new study seemed to offer a new range of dependent variables, for instance the use of computers in classrooms and achievement scores for those studying French and English as a foreign language, mathematics, science, civics, reading, and writing (IEA, 2001, 1993, 1992a, 1992b, 1976, 1975, 1973).

The informality of IEA as an organization proved to be a significant obstacle to the carrying out of each survey. From the first meeting in Chicago in the mid–1950s up to the publication of the Second International Study of Science (SIMS) in 1982, IEA had remained an association of academic institutes unaffiliated with governmental agencies. In centralized systems, such as in France and Japan, the member institutes happened to be sponsored by governments. But in the United States, as well as in many low- and middle-income countries, the representative to the IEA was an independent professor with no official capacity.[2] After the 1982 results suggested that the United States was behind other industrial democracies, the U.S. Department of Education was called to account by Congress, only to discover that the U.S. samples for SIMS were not nationally representative. The National Center for Education Statistics (NCES) and the National Science Foundation concluded that for political as well as professional reasons, this lack of representativeness should not re-occur. Any international study in which the United States participates (particularly with such visibility) henceforth must insure that the data meet normal standards of representativeness, reliability, and validity. Thus, the timely delivery of the IEA Third International Mathematics and Science Study (TIMSS) became a high federal priority. Coordinated with the stated U.S. goals of being first in math and science (Vinovskis, 1999), the TIMSS study was the most ambitious and expensive enterprise undertaken by the IEA, and its results have generated debate that should extend well into the future (Beatty, 1997; Peak, 1997, 1996; Ravitch, 2003; Baker and LeTendre, 2005). But how could the federal government enforce its standards upon a private voluntary organization of which it was not even an official member?

THE ORIGIN AND PURPOSES OF BICSE

In 1988 two agencies (the National Science Foundation and the U.S. Department of Education) decided to sponsor the Board on International Comparative Studies Education (BICSE) through the National Academy of Sciences. The board was given four main objectives: (1) to suggest technical principles which could be internationally enforced; (2) to guide U.S. agencies on the types of studies of highest priority;[3] (3) to ensure the efficient management of any studies in which U.S. students might be studied; and (4) to sensibly coordinate the necessary financing and management of the infrastructure necessary to make international studies viable over the long run. The BICSE committee was influential globally over the next dozen years. While many nations did not feel comfortable working with the U.S. government in education,[4] they listened more attentively to the views of a committee of the National Academy of Sciences because it was perceived to be professionally competent and politically neutral.

BICSE's first report, distributed in 1990, consists of a series of reasons why a country might wish to engage in international comparative work in education and the kinds of studies that might be useful. The report includes rationales for studies of academic achievement as well as a wide range of educational questions using descriptive, qualitative, observational techniques (Bradburn and Gilford, 1990). The eclectic nature of the list helped overcome some of the misgivings of the

academic community. The report laid out a series of principles for international standards in areas such as technical validity, research neutrality, sampling, access to schools, instrument construction and administration, analysis, reporting, and dissemination. Unlike textbook and other academic sources for these standards, the BICSE report had a scientific imprimatur which influenced public agencies financing U.S. participation in international studies, agencies in the participating countries cooperating with the United States, and inter-governmental agencies (such as OECD, UNESCO, and the World Bank).

The next BICSE report, three years later, went well beyond the first and outlined the international structures required to put a larger and more professional agenda of studies into effect (Gilford, 1993). It included sections dealing with problems of comparability across nations, financing requirements, and suggestions for improving the international infrastructure for cross national research. As a result of these first two reports and the multiple informal discussions surrounding them, IEA was transformed from an informal 'academic club' to a professional multi-national agency. It initiated and enforced a dues structure, technical procedures for implementing studies, and formal sanctions for participating countries which did not adhere to them.[5] BICSE published additional reports on the range of international studies available (National Research Council, 1995), how to get the most out of re-analyses (National Research Council, 1999) and how to use cross national studies in the making of educational policy (National Research Council, 2002).

However successful BICSE was in establishing a consensus among nations, cross-national studies continued to be challenged for many reasons. Some believed that the United States was at a systematic disadvantage because it had higher enrollment rates and comprehensive schooling, while other participating countries with lower enrollment rates and elitist systems (and thus unrepresentative samples) enjoyed artificially high scores (Rotberg, 1990). This stimulated a powerful, personal reply from some BICSE committee members (Bradburn, Haertel, et al., 1991).

THE ORIGIN OF OECD AS A LEADER IN EDUCATION STATISTICS

The U.S. federal government became a strong supporter of cross national education research following the publication of *A Nation at Risk* (National Commission on Excellence in Education, 1983). Among other things, the report claimed that international comparisons indicated that the United States had surrendered in the battle to create the most knowledgeable citizenry in the world. The U.S. government decided to utilize the Organization for Economic Cooperation and Development (OECD) to spearhead an initiative to improve international education statistics.[6]

An acrimonious meeting took place one year after the publication of a *Nation at Risk.* The U.S. delegate from the Center for Educational Research and Innovation (CERI) board challenged OECD to undertake a project to quantify educational quality with measures for curriculum standards, costs, sources of financing, learning achievement, and employment trends.[7] The reaction among the OECD staff at that time was one of shock and suspicion that quantification of educational quality and especially cross national testing constituted a right-wing political tactic, and an abrogation of professional standards.[8] Like Piaget in 1933, they believed that education was specific to each culture, and therefore could not be measured between cultures. They argued that the process of quantification would oversimplify and misrepresent a nation's education system. Many assumed that the demand for such information would shift as soon as the political party of the U.S. presidency changed.[9] It is now clear that this was not the case. In the 20-year interim, the OECD helped pioneer the International Education Indicators project which led to an annual report titled *Education at a Glance*—the most highly demanded of all OECD publications (OECD, 2002).[10]

The Origin of IAEP, PISA, and IALS

Another important topic of discussion at the BICSE meetings concerned whether the U.S. should participate (and help generate) curriculum-based international tests, or whether they should pioneer a new international initiative concentrating largely on a criterion-based achievement test variable roughly analogous to the National Assessment of Educational Progress (NAEP). The original IEA studies which had generated a myriad of explanatory variables were thought to be superior in the sense that they could offer a wide variety of hypotheses.[11] But the IEA style of studies required a decade to generate, were extraordinarily complex to implement, and were expensive. Policy makers said they wanted a faster turn-around with data, including trend data, delivered on time, and collected more frequently.[12]

The U.S. Department of Education agreed to a trial of a NAEP-like international project utilizing the same center that had been administering the NAEP at the Education Testing Service. A new project, called the International Assessment of Educational Progress (IAEP), saw to it that instruments were developed and data collected, analyzed, and reported for more than a dozen countries, within a record time of three years. The measure of academic achievement was developed centrally and based largely on what a student was supposed to perform at a given age level instead of being developed from an analysis of each country's curriculum (IAEP, 1992, 1991).

Many in the international education community felt that the IAEP project was misguided in its decision to determine, without sufficient analysis of local curricula, what a student should be able to do. Although BICSE recommended that the United States participate in both styles of analysis, it was not clear whether the cost of double participation could be justified. However, the timeliness, credibility, and immediate demand for information from the international indicator project helped insure that it would have a future. Since the IEA seemed reluctant to engage in an enterprise that monitored achievement without studying its antecedents in depth, the task for developing the new project was assigned to OECD. It would later develop into the Program of International Student Assessment (PISA) (OECD, 2004a).

PISA has proven to be a wide-ranging and valuable assessment. Among its chief innovations is the introduction of a cross-curricular problem-solving component. It examines students' capacity "to understand problems situated in novel and cross-curricular settings, to identify relevant information or constraints, to represent possible alternatives or solution paths, to develop solution strategies, and to solve problems and communicate the solutions" (OECD, 2004b, p. 3). This component helps balance the heavy weight previous assessments placed on memorized knowledge. Yet interestingly, the rankings in PISA were largely similar to those in previous surveys. Reading ability was divided into five levels of proficiency. Of the 39 countries participating in the PISA project, only 5 percent of the 13-year-old school children in Peru could read above the international mean whereas 75 percent of the 13 year old students in Korea could read above the international mean. The United States, with 60 percent of its 13-year-old students above the mean, was ranked 15th (Figure 7.2.).

OECD was also encouraged to develop a new survey to gauge adult literacy, newly defined as a performance indicator.[13] The results of the International Adult Literacy Study (IALS) were startling. Several nations, such as Poland, which had heretofore been known for having very effective educational systems, had pockets of genuine illiteracy among its adult population and a generally low level of literate performance relative to what had been expected. Only 2.9 percent of the adults surveyed in Poland performed at Level 4 literacy.[14] The United States, which lags behind many counties in elementary and high school surveys, had 17.3 percent of its adult population performing at Level 4, compared to only 12.3 percent in Germany, 14.6 percent in Netherlands, roughly 9 percent in Switzerland (OECD, 2005). Furthermore, no country except

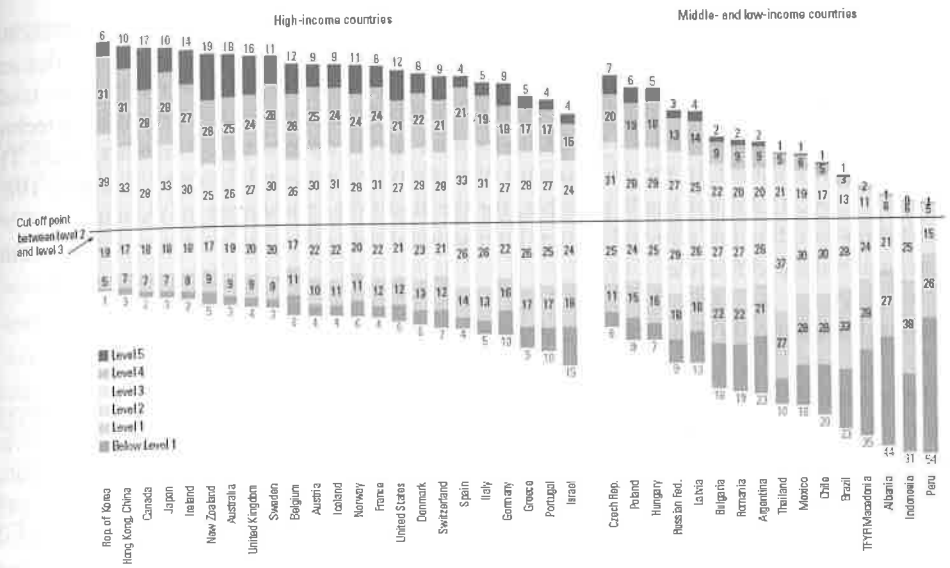

FIGURE 7.2 Percentage of 15-Year-Old Students in Five Proficiency Levels of Reading, 2000–2002.
Source: UNESCO (2005b), *EFA Global Monitoring Report 2005: The Quality Imperative*. Montreal: UNESCO.

Sweden had more than 5 percent of its adults surveyed perform at Level 5. These findings added new urgency to debates concerning adult education across the globe.

Origin of Date on Education Expenditures

Descriptive data on education expenditures was also an important product of the effort to improve education statistics (Sherman, 1982, 1981). The first education finance comparability study was completed for ten countries in 1995 and twenty countries in 1999. After a series of improvements in design and clarity of definitions, the survey was updated using trend data in 2002 and 2004. Today, an expert group governs the collection and reporting of education finance data at OECD.[15] They focus on data whose quality is relatively consistent across geographical regions and which can be collected in a timely manner (Charbonnier, 2005a).

Six financial indicators meet these criteria: (1) expenditures/student; (2) expenditures on primary, secondary, and tertiary educational institutions relative to GDP; (3) relative proportions of the public and private investment in educational institutions; (4) total public expenditures on education; (5) support for students and households through public subsidies; and (6) expenditures on institutions by service category (technical, general, secondary, etc.) and by resource category (public and private). Three new indicators are under development—a metric combining expenditure and performance, an empirical listing of who pays and who benefits from university education, and an indicator of cost/university graduate (Charbonnier, 2005a).

Explanation for the Success of the OECD Project

One key to the success of the OECD effort has been the way in which quality was controlled. Instead of being engineered by a single central authority, quality, and efficiency of implementation was established on the principle of informed peer pressure. Because each nation's reputations

depended on the findings, each nation worked hard to make sure that the results meet commonly accepted standards of rigor. All partners have equally enforced adherence to time schedules and quality control. This distributed control allowed the OECD to remain comfortably in its traditional role of neutral coordinator of the interests of member countries while still ensuring technical precision. It also has allowed the OECD to explore terrain opened up by member states. The United States, for instance, in the sponsorship of work on educational finance (Sherman, 1982, 1981; Barro, 1998, 1996) has made possible, for the first time, consensus on technical matters. This consensus-building has allowed regular international collection and analysis of data on education expenditures and finance.

The Use of International Education Statistics

Since the beginning, public officials have often used international studies of academic achievement not as an opportunity to better understand the nature of learning, but as a means of motivating internal reforms in the name of global competition. The results of the Second International Science Study (SIMS) in the 1980s showed that the United States was at best a mediocre performer. Japanese students were reported to learn over 60 percent of what they were taught while the American students learned only 40 percent (Baker, 1993, p. 19). Findings that the U.S. scores were low by comparison to other countries were pervasive and consistent, causing some to suspect that the data might have been mishandled by the federal government so as to generate greater support for educational reform (Bracey, 1996),[16] a hypothesis that was rejected by others (Baker, 1997). Some wondered if it was fair to compare a large and complex nation such as the United States with smaller more homogeneous nations. If states are treated as nations,[17] five of the ten highest scoring nations in 13-year-old math proficiency were U.S. states. These included Iowa (ranked 2nd in the world), North Dakota (3rd), Minnesota (5th), Maine (8th), and New Hampshire (9th) (National Center for Education Statistics, 1996, p. 155).

BICSE helped organize and ensure the timely delivery of the IEA Third International Mathematics and Science Study (TIMSS). But when results appeared in 1995, the United States was again found to be a low achiever. In terms of increases in math between the 4th and the 8th grades, the U.S. gain (93) was significantly lower than Thailand (168); in this the United States ranked last among the 24 nations in the sample (Martin, Gregory, and Stemler, 1999). Much of the problem of U.S. scores was attributed to the breadth of the curriculum, which allegedly precluded the necessary depth (Schmidt, McKnight, and Raizen, 1997). However more recent analyses suggest that national achievement scores in math and science are only vaguely related to characteristics of the curriculum (Baker and LeTendre, 2005, p. 162).

With the follow up administrations of TIMSS in 1999 and 2003 (and one expected in 2007), the U.S. placement may be becoming more difficult to interpret (Martin, Gregory and Stemler, 1999). Eighth grade scores in math and science were higher in 2003 than in 1995 relative to the other 21 countries. However, fourth-grade scores were lower in 2003 than in 1995 compared to other countries (Gonzales, Guzman et. al., 2004). New, more comprehensive data have allowed the posing of new questions and concerns. Instead of using average achievement scores across nations, now it is common to ask which nation is more successful at closing gaps in educational disadvantage among students.[18] In this the United States ranks 18th among the 24 wealthy industrial democracies (UNICEF, 2002).

Does an increase in education spending increase children's math and science performance? Since the work of James S. Coleman and others (1966) and Christopher Jencks (1972) in the United States, it has been common to conclude that schools, especially public schools, are relatively ineffective and inefficient at raising their levels of performance. Drawing on new interna-

Country	Change in Math and Science Score 1970–1994	Increase in Real Spending Per Pupil 1970–1994	Increase in Real GDP/Capita 1970–1994
Austria	–2.3	269.8	46.4
New Zealand	–9.7	222.5	24.3
France	–6.0	211.6	60.7
Italy	1.3	125.7	74.6
Germany	–4.8	108.1	66.8
Japan	–1.9	103.3	100.7
United Kingdom	–8.2	76.7	58.3
Belgium	–4.7	64.7	68.0
Netherlands	1.7	36.3	52.9
United States	0	33.1	70.5
Sweden	4.3	28.5	35.1

FIGURE 7.3 Percentage Changes in Test Scores and Real Expenditures Per Pupil 1970-1994.
Source: UNESCO (2004), *2005 EFA Monitoring Report: The Quality Imperative*. Paris: UNESCO, p. 65.

tional data available in the 1970s, one study across 29 countries found that the explanatory power of children's social background varied by country, and in poorer countries was exceeded by the explanatory power of school quality (Heyneman and Loxley, 1983). These results suggested that when highly motivated, children from impoverished backgrounds can use the school system to overcome the exigencies of their social status.

The availability of additional cross national data in the late 1990s allowed several new analyses (Harris, 2007; Baker, Goesling, and LeTendre, 2003; Gameron and Long, 2006; Hayneman, 2005). These new analyses suggest that the "Heyneman-Loxley Effect" (low impact of social background and high impact of school quality in low-income countries and the reverse in high-income countries) continues. With respect to the modest impact of new resources on achievement (e.g., between 1970–1994), Austria increased real spending per pupil by 269 percent, but its math and science scores declined by 2.3 percent over the same time period. New Zealand increased its spending per pupil by 222 percent, yet its math and science scores declined by 9.7 percent (Figure 7.3.).

Estimates of the impact of school resources on achievement tend to be greater in developing countries than in the United States. For instance, teacher education is a positive and statistically significant predictor of student performance in 56 percent of the estimates for developing countries but in only 9 percent of those in the United States (see Figure 7.4). School resources is a positive and statistically significant predictor of math and science achievement in 65 percent of the studies conducted in developing countries but in only 9 percent of those conducted in the United States.[19]

The existence of new, high quality, international data sets continues to generate innovative hypotheses about the nature of American schooling. Some have combined the different data sets from school achievement (TIMSS and PISA) with adult literacy (IALS) and have begun to ask questions of new kinds. Might differing forms of social segregation in education (between and within schools and classrooms) be related to social gaps in adult literacy? Using social segregation within school systems, for instance, countries seem to fall into groups. The Nordic and East Asian nations have the lowest levels of social inequality in schools; English-speaking nations

	% Positive and Significant	
	United States[a]	Developing Countries[b]
Real Classroom Resources		
Pupil/Teacher Ratio (PTR)	14	27
Teacher Education	9	56
Teacher Experience	29	35
Financial Resources		
Teacher Salaries	20	31
Expenditures Per Pupil	27	50
Other		
Facilities	9	65

a. Based on 376 production function estimates.
b. Based on 96 production function estimates.

FIGURE 7.4 Effect of Key Resources on Student Performance.

Source: UNESCO (2004), *2005 EFA Monitoring Report: The Quality Imperative*. Paris: UNESCO, p. 65.

have a much greater level and German-speaking countries (Germany and Austria) have the highest level. Social segregation within school systems is associated with differences in adult literacy. Perhaps Germany's practice of segregating students into secondary schools with terminal functions (rare in Europe) helps explain the high levels of inequality in German adult literary (Green, Preston, and Janmaat, 2006, p. 125).

Some researchers have combined new international data sets on achievement with other sources of information on school curriculum and textbooks (Heyneman, 2006a). Mathematics textbooks and curricula today cover more topics and in more depth (Cummings, Knipe et. al., 2007). For many years, schools have emphasized higher order cognitive skills. Now it is being suggested that an emphasis on higher order cognitive skills has begun to boost average IQ scores. This suggests that the power of schools and of school systems to affect human behavior may have been underestimated (Cummings et. al., 2007).

Perhaps the biggest puzzle is not whether these new theories of school performance in an international context are necessarily correct, but rather how they have come to exist at all. How did the world go from a debate in 1933 over whether one could even count schools to a debate over the influences of schools into adulthood and across generations? The existence of new and improved statistics has facilitated new questions and new theories, and has justified the effort of the agencies and the individuals who have been involved.

EDUCATION OUTSIDE OF THE OECD COUNTRIES

Of the pupils enrolled in schools around the world, only 12 percent are enrolled in OECD countries. If high quality educational statistics were to be confined to OECD countries, 88 percent of the pupils and schools in the world would be left out.

The Role of UNESCO

Country-specific descriptive educational statistics are the responsibility of each Ministry of Education. The office of statistics in UNESCO circulates a template with instructions on their collec-

tion. It includes common definitions and suggests standards for quality control. These statistics are then published annually. The International Institute of Education Planning, a sub-unit of UNESCO, is responsible for training public officials in education planning including the application of UNESCO statistical templates. UNESCO receives data from each member state, and then has the responsibility to determine its accuracy. If the data appear to be below an acceptable standard, or if there are important questions about the data, UNESCO will inquire about its origins. In instances where corrections cannot be made or when the data appear to be particularly implausible, UNESCO reserves the right to not publish a country's statistics in its annual statistical yearbook.[20]

During the debt crises in the 1980s, the quality of schooling in low and in many middle income countries declined (Heyneman and Fuller, 1989).[21] The quality of education statistics also declined. Figures on rates of enrollment and on numbers of teachers became increasingly untrustworthy. The statistics division within UNESCO faced declining budgetary support and increasing difficulties monitoring the quality of statistics from low-income countries. When statistics of questionable quality were received, fewer professionals were available to intervene. The declining quality of education statistics began to hamper the ability of nations to accurately review their own priorities and the ability of international organizations to provide assistance (Heyneman, 1998, 1997a, 1997b). The mounting inadequacies in educational statistics stymied international cooperation by making it difficult to accurately identify the sources of problems and the means of amelioration (Heyneman, 1999, 1993a).

BICSE was constituted primarily to improve the comparability of the United States with other industrial democracies. But BICSE members became increasingly aware that American trade and social interests included Brazil, Russia, China, Indonesia, India, and many other countries which were not OECD members. Hence, it became necessary to extend the interests of BICSE to improving the quality of statistics in non-OECD countries. In an effort to become informed as to the current status of education statistics in UNESCO, a BICSE consultant (supported by the World Bank) argued that the decline in the quality of education statistics was a serious problem even to the United States (Puryear, 1995; Heyneman, 2003a).

The U.S. government was not a member of UNESCO, but it made informal contact,[22] and through BICSE, helped sponsor a specialized assessment of UNESCO's statistics capacity. The study was highly critical of UNESCO's handling of statistics, but was also careful to set out feasible organizational options to improve the situation (Guthrie and Hansen, 1995). These options were adopted by a special UNESCO commission chaired by Jo Ritzen, the Minister of Education from the Netherlands. The division of education statistics was reconfigured to become an independent institute, chaired by the ex-president of the British Statistical Association, staffed with newly acquired professionals, and relocated from Paris to Montreal. The improvement in UNESCO's statistics capacity in the interim has been significant (Brown and Micklewright, 2004; UNESCO, 2005a).

Despite this marked improvement in the professionalism of its publications (see for instance UNESCO, 2005a), the reconfiguration has not gone to the root of the problem, which is the poor state of education statistics generated by many of the countries themselves. Of the 196 countries reporting, 30 percent (59) did not have acceptable data on enrollment as recently as 1999. This includes 40 percent (18) of countries in Sub-Saharan Africa. When independent household surveys were employed to counter-check the accuracy of government enrollment statistics in Sub Saharan Africa 68 percent (17) were found to have significantly lower enrollment rates than were submitted, causing UNESCO to questions "the reality of enrollment statistics in a good number of countries" (UNESCO, 2002, p. 49).

Improving education statistics in low-income countries cannot be achieved through UNESCO

alone. Precedent had been set by the speedy and effective collaboration within OECD countries. This experience might be useful more broadly.[23]

In 1996, the World Bank helped finance a program of collaboration between OECD and UNESCO. The result, called the World Education Indicators (WEI) Project, allowed non-OECD countries to develop indicators comparable to those of OECD countries. Many of the countries most eager to collaborate (Russia, China, Pakistan, India, Brazil, and Indonesia) were of such size that in 2002, the OECD publication *Education at a Glance* could claim to cover roughly two thirds of the world's population (OECD, 2002).

In the 1970s and 1980s, IEA surveys of academic achievement were greeted with skepticism by those in the international education community who worried that middle and low-income countries would not benefit from an academic Olympics. With the broadening of objectives and insights, and the creation of regional surveys of achievement, their hesitancy has been modified. Publications such as *Sub-Saharan Africa Regional Report* (UNESCO, 2001b) and *Latin America and the Caribbean Regional Report* (UNESCO, 2001a), and similar reports from the Southern and Eastern African Consortium for Monitoring Educational Quality (SACMEQ) have proved useful in assessing countries with financial and cultural similarities (Heyneman, 1993b). On reflection today, educational leaders from developing countries can see the advantages of having participated in international surveys of academic achievement (Brunner, 2005). In the future, an increasing variety of countries will likely participate in these surveys. International agencies are likely to continue to sponsor their regular participation.

Education-For-All

Among the more important international objectives has been the achievement of universal enrollment at least in primary school. Since the end of WWII, universal primary enrollment has been achieved in 85 countries, including most of Latin America, Europe and Central Asia, much of East Asia, and Iraq, Tunisia, Jordan, Algeria, Botswana, Gabon, Lesotho, South Africa, Swaziland, Togo, and Uganda (UNESCO, 2002, p. 46). Nine nations remain with gross enrollment rates below 70 percent, although there are many additional countries where recent data are not sufficient to determine enrollment rates. Despite significant progress, most countries fall short of the standard of universal completion of grade nine. In West and Central Africa and in South Asia about one third of the deficit in completion rates can be attributed to dropouts; in Central America and in Eastern and Southern Africa, dropouts constitute 70 percent of the missing completers; in Europe and Central Asia and in Latin American dropouts constitute 80 percent and 92 percent, respectively (Pritchett, 2004, p. 181). The remainder is accounted for by those who repeat.

The main problem is not lack of access to primary education but progress through the grades once access has been obtained. While some national school systems have been found to operate efficiently despite low levels of resources (Heyneman, 2004a), it is also the case that low resources often constitute the largest explanation for low achievement (Heyneman and Loxley, 1983). Low achievement is one cause of dropping out of school, and the differences from one school to another in available resources may be one of the principal causes of low achievement (Heyneman, 2004a).

Using newly available financial data from UNESCO, we have taken the level of expenditures on education as a percentage of gross domestic product (GDP) and the allocation to education (in $US). We divided the population of school age children by educational expenditures to find the expenditures for each school age person, and have ranged the countries by their place on a world spectrum (Figure 7.5).

Low-income countries are able to allocate to education only a small fraction of what a high-

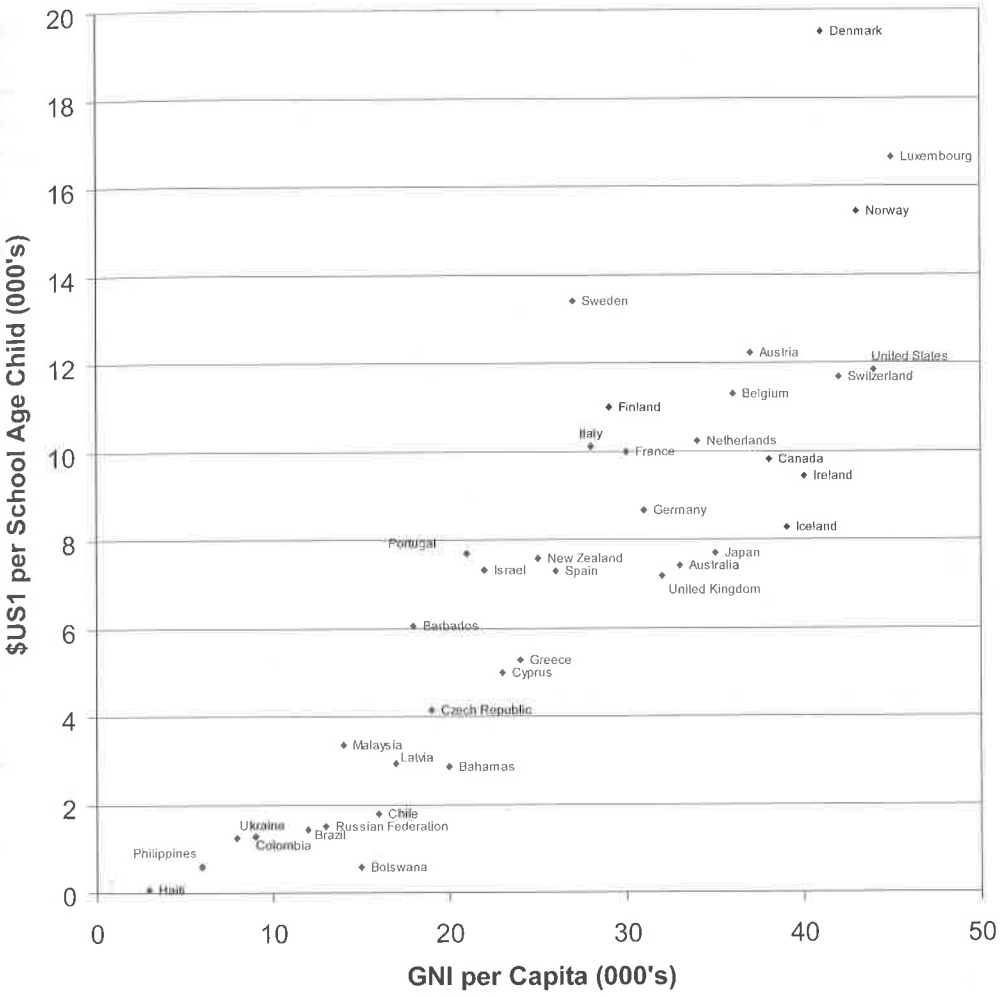

FIGURE 7.5 Equity of International Educational Resources.

income country might allocate. Bangladesh spends about $157/school age person/year. This compares unfavorably with Azerbaijan ($486), Brazil ($1450), and the Netherlands ($19,468). To mitigate this widespread inequality, the international community has declared that completion of basic education at a reasonable standard of quality is a fundamental right for all children. At a summit in Thailand in 1990, Education-For-All was adopted as a global priority.[24] Development assistance agencies and countries themselves were expected to allocate more resources to education and a higher percentage of education resources to basic education. Early in the 1990s, however, many scholars realized that the statistical capacity was too weak to monitor changes in basic education and that many of the Education-For-All objectives, such as achieving basic levels of literacy and numeracy, were simply impossible to prove (Bloom, 2005). This lack of direct proof stimulated UNESCO to create proxy measures of education progress (UNESCO, 2004, Appendix 1, p. 236).

Yet despite a weak statistical capacity, it was not a secret that the targets established in 1990 would not be reached (UNESCO, 2004, 2002; World Bank, 2005), and that new aims would have to be given. Debate ensued over how much Universal Basic Education would cost. UNESCO

estimated that it would cost $US 4.3 billion (UNESCO, 2004, p. 136). UNICEF estimated that it would cost $US 4.9 billion (Heyneman, 2004c). The World Bank has estimated that it would cost $US 8.4 billion (World Bank, 2005). The difference among the estimates reflects differing positions regarding minimum per student expenditures, whether enrollment or completion is taken as the objective, whether only the 47 poorest or all developing countries are included, and whether special funds are set aside for AIDS orphans (Glewwe and Zhao, 2006).

The barrier to achieving Education-for-All objectives may have little to do with its cost and more to do with the political will of the participants. The portion of foreign aid allocated to education has not changed in two decades. It accounts for 6 percent of new commitments in USAID and 8 percent in the World Bank (Heyneman, 2006a). Some scholars suggest that the recipient countries have not reallocated public expenditures toward education in part because the opportunities for rent-seeking are fewer in education than in other sectors with more capital expenditures (Corrales, 2006).[25]

Role of the World Bank

Two factors have limited the World Bank since it began education lending. One was the use of manpower forecasting, which held an analytic monopoly between 1962 and 1980. Because justifications for public expenditures could only be made for vocational and technical education, the Bank confined lending to those areas. Primary education was largely excluded,[26] as was lending to improve the arts and sciences in tertiary education. Secondary education could be justified if the schools could be equipped with metal shops, wood shops and domestic science laboratories for girls. Although adding these facilities more than doubled the unit expenditures, these practical skills, it was argued, were necessary to offset the overemphasis on academic subjects which, according to Bank staff, led to educated unemployment (Heyneman, 2003b). The second handicap was the relative absence of experienced education staff. To alleviate this latter problem in 1977, the World Bank established the Education Finance Division in UNESCO with staff assigned to work on World Bank operations.[27]

Disagreements quickly emerged over issues of sector policy. As the designated education agency within the UN system, UNESCO argued that it alone should be responsible for education policy. Representatives to UNESCO were ministers of education. The Bank worried that UNESCO would advocate for education without sufficient emphasis on cost containment or efficiency. Representatives to the World Bank were ministers of finance. The Bank believed that its direct link to the financial ministries placed it in a better position to make economically feasible education policy. In the 1970s, the Bank began to acquire its own education policy staff and establish its own education sector policy.

Bank staff also came to recognize that the monopoly of manpower forecasting was creating distortions in the lending program. As a remedy, the Bank hired staff who understood the economic technique of rate-of-return analysis so that it could begin to integrate economic rates of return into its education sector work. Because this new technique utilized changes in monetary productivity, the benefits of education could be calculated for any level including primary education and not be confined to vocational or technical education. The changes in the Bank's methodologies, announced in the Education Policy Paper of 1980, allowed the Bank for the first time to include assistance to all levels of education and, theoretically, to any educational purpose whether vocational or academic. As a result, the Bank's analytic work led it to become a leader in the Rates of Return to Education studies.

The Bank's analyses have consistently found that the *public* Rates of Return to Education tend to be greatest for investments made in primary rather than higher education and for low-

income countries (Psacharopoulos, Tan, and Jimenez, 1986; Psacharopoulos, 1994). This was the case in part because of the very low levels of investment in the quality of primary education. Based on these findings, the Bank began recommending that countries shift public expenditures from higher education to elementary education; that the private share of the costs for higher education should be increased; and that loans should be made available for those who cannot afford the newly increased fees (World Bank, 1995, 1994).

Many borrowing countries, particularly in Sub Saharan Africa, objected to the suggestion that they reduce public expenditures on higher education on the grounds that they would not be able to develop expertise in vital areas such as health, agricultural engineering, and general science, and would end up being continuously dependent upon the advice of those who share neither their culture nor the national priorities. As the use of RORE grew in importance, some researchers objected to the Bank's interpretations.[28] They argued that while there were numerous problems in the RORE calculations, the real issues lay in how the results were interpreted and how recommendations were made to countries. Low public returns to higher education might imply that the efficiency or quality of higher education should be improved. Instead of suggesting that the solution lay in a reduction of public expenditures, the solution may well lie in an increase in the effectiveness of public expenditures (Heyneman, 1995).[29] Some argued that economic returns are an ancillary result of education, and that the true returns come in other forms, such as social cohesion (Merrett, 1971; Heyneman, 2000).

These criticisms began to call into question the credibility of the Bank's analytic staff. In response, the Bank contracted an external panel of higher education experts to write a new report on higher education. This report concluded that the earlier interpretations had seriously underestimated the non-monetary benefits of higher education (Task Force on Higher Education and Society, 2000, 2004). Following the recommendations of the Task Force, the Bank began to incorporate non-monetary rationales into the deliberations over lending strategies, and now offers RORE as one among many sources of information as to how best to invest in education. The role or RORE in the Bank's lending policy is now to highlight potential problems, rather than to rigidly define priorities (Heyneman, 2003b).

The World Bank's owners insist on policy conditionality; hence conditioning lending on an efficient allocation of public expenditures is a normal Bank responsibility.[30] Policy conditionality may include environmental safeguards for a transport project or merit pay in an education project. Technical staff of the Bank believe the conditions are necessary for a successful investment. The problem is that investments have failed because the conditions, on hindsight, were based on false premises (Bennell, 1996; Colclough, 1996; Heyneman, 2003b, 2006b).

SUMMARY

Education finance would be remiss to forget either Piaget's caution or Dottrens' confidence toward international comparisons in education. Piaget was correct when he outlined the several obstacles to obtaining useful, comparable knowledge about education systems. But looking back, it is easy to see that his list of "obstacles" has functioned as a catalogue of challenges rather than a series of insurmountable impasses. The first of these accomplishments was the development of the International Standard Classification of Education (ISCED) in 1951, which ensured that all members of the international community had a commensurable set of definitions. With this conceptual groundwork laid, the next challenge was to create indicators which gave the clearest picture of, to paraphrase Dottrens, "what other countries were doing." In the 1970s, significant advancements were made in this area—for the first time, reliable measurements of education

achievement and outcomes (and not just inputs) were included in international comparisons.

Since the 1970s, the refinement of measurements of outcomes has become the dominant interest in international comparative education. While the United States and other OECD countries have the longest history of measuring and comparing themselves, there has been a surge in participation by non-OECD countries as well. Much of this positive trend has been made possible by the influx of foreign aid. Yet, as Dollar and Pritchett show (1998), a disturbing portion of foreign aid has failed to yield significant improvements in recipient nations' ability to monitor their own education systems. The aspirations motivating Education-For-All offer hope that future findings will serve the international community just as well as Piaget's caution—as a list of goals to accomplish, rather than a series of excuses to abandon the comparative project.

Debates about the validity and purposes of cross-national statistics are a sign of a healthy research community. Governments, traders, and humanitarians will continue to pressure scholars to expand the breadth and depth of statistical projects throughout the world. If the past is any indicator of the future, the new theories and insights produced by this research will justify these efforts.

Important progress in the quality of education statistics has been made outside of OECD countries. The former Soviet Union, the countries of eastern and central Europe, and the emerging democracies in Latin American and Asia have demanded, and often achieved, equal partnership in cross-national educational exercises. Nevertheless, many important areas of the world have made little progress in their ability to gather and analyze important information about their own educational systems. The future dilemma is stark: will educational progress be left unmeasured or will these nations be supplied with the resources and the regulatory environment to join others in their cross-national efforts to monitor educational progress?

NOTES

1. Dottrens and Piaget were Swiss educational psychologists. Dotterns served on the Board of Directors of the International Bureau of Education (IBE); Piaget served as IBE's director for 40 years. The International Bureau of Education was founded in Geneva in 1925 as a non-governmental educational organization. In 1929 it allowed countries to join as members. Today, the IBE is a specialized agency within UNESCO.
2. The United States was represented on the IEA General Assembly by Richard Wolff of Teachers College, Columbia University.
3. There were three main concerns: the use of (1) academic achievement versus other dependent variables; (2) curriculum-based (IEA) studies of achievement versus census-based (ETS and OECD) studies of achievement; and (3) empirical as opposed to qualitative (video-taping) models of schooling.
4. Reasons for reluctance varied. They included a feeling of imposition from a well-funded dominant authority and a tendency to associate some educational policies which encourage competition with neo-liberal political views.
5. The most serious sanction was for a country to be left off of the official report, and allocated space "below the line" indicating that there were sufficient technical reasons to declare their results non-comparable. This sanction was enforced, for instance, in cases where sampling was not representative of the country.
6. The United States had withdrawn from UNESCO, the second international agency with an education statistics mandate.
7. The CERI is an advisory body for the education and labor branch of the OECD.
8. Personal communication with OECD officials in 1984.
9. When assessing U.S. policy trends, one common mistake of Europeans is to underestimate the importance of local officials in setting the national education agenda. The extent to which the U.S. president

can, without the support of local officials, set a national educational agenda, is less than in many European countries.

10. The publication *Education at a Glance* reports annually on trends in enrollment, finance, and changes in management structures across OECD countries.
11. From the beginning the IEA-sponsored studies were controlled by academics, not policy makers, hence IEA project managers felt that they could afford the time to gather information on a wide variety of influences on academic achievement. These included separate instruments for school principals, students and teachers; a separate analysis of school curriculum and (occasionally) cases studies of classroom pedagogy in each participating country.
12. Clare Burstall, Director of the National Foundation for Education Research in England and Wales (1983–1993), made this case to a meeting of the BICSE Board in Washington, D.C.
13. The UNESCO definition of literacy was vague and applied with unpredictable variation. As a result, for purposes of literacy statistics, the world had been divided into two groups: OECD and non-OECD (Heyneman, 1998).
14. Literacy was divided into five levels of proficiency, and countries were assessed by what percentage of the population had achieved levels five, four or three (the three top proficiency levels).
15. Current work includes the clarification of systems of pensions, student loans, interest payments, research and development and ancillary services (Charbonnier, 2005b).
16. Both scholars and politicians, however, had to acknowledge that the performance of American students relative to students in other industrialized democracies was considerably better in reading than it was in math and science (Elley, 1992).
17. This is not an irrational comparison. Minnesota for instance, has a population equivalent to Norway.
18. Educational disadvantage is a compendium of the percentage of children scoring below a fixed international benchmark in surveys of reading literacy of 15-year-olds, math and science literacy of 15-year-olds, and math and science eighth-grade achievement.
19. These comparisons do not account for the possibility that there may be variations in the quality of the studies.
20. This "right of refusal" is used by all UN statistical agencies with mixed success. When the Soviet Union dissolved, each of the 15 republics argued that their own definitions and calculations for current accounts and currency reserves were superior to those of the IMF. The IMF simply left blank the figures which were not calculated in accordance with their standards. Within a few years, all 15 republics had accepted the IMF guidelines. Leaving figures for a nation's education enrollment blank may not create the same problem as leaving blank the nation's statistics on current accounts or currency reserves.
21. In Latin America the 1980s are referred to as the "lost decade."
22. With expenses covered by the World Bank, Jeanne Griffith from the National Center for Education Statistics met with senior officials of UNESCO in Paris in the winter of 1994.
23. It is more complex and more difficult to achieve collaboration between international agencies than it is to achieve collaboration between different agencies, for instance, of the U.S. federal government.
24. The Education-For-All initiative was sponsored by UNESCO, UNDP, UNICEF, and the World Bank. The latter was a reluctant participant. Similar global conferences had been used to force countries and multilateral lending institutions into making financial commitments to special interests without consultation with fiscal authorities. In the case of Education-For-All, agreements were signed by Ministers of Education who have no fiscal authority. Hence, commitments involving a reallocation of public resources were null and void from the beginning. Also, see chapter by Fiske and Ladd in this volume.
25. "Rent seeking" in this case refers to the opportunity to use a government office to collect illegal fees or bribes ("rents"). It is argued that the opportunity to collect rents is higher in capital intensive sectors such as infrastructure rather than in the labor-intensive sectors such as education.
26. Exceptions were made for community primary schools equipped with facilities for community meetings and other facilities for adult education.
27. A similar cooperative program was established by the World Bank in the Food and Agricultural Organization (FAO).

28. See for instance: Barnett and Finnemore (1999), Bennell (1996), Colclough (1996), International Labor Office (1996), Carnoy (1995), Hinchliffe (1993), Biersteker (1992), Helleiner (1992), Colclough and Manor (1991), and Merrett (1971).
29. While public returns to investments in higher education were commonly lower than public returns to primary education, it was also common to find that private returns to higher education were lower than public returns because of the high public subsidies in higher education. In fact, higher education was often free of tuition whereas in primary education tuition was common.
30. The international development banks are owned by its member states. Countries own differing shares, much like a private equity company.

REFERENCES

Baker, D. P., Goesling, B., & LeTendre, G. (2003). Socioeconomic Status, School Quality and National Economic Development: A Cross-National Analysis of the 'Heyneman-Loxley' Effect. *Comparative Education Review, 46*(3), 291–313.

Baker, D. P., & LeTendre, G. (2005). *National Differences, Global Similarities: World Culture and the Future of Schooling*. Stanford, CA: Stanford University Press.

Baker, D. P. (1993). Compared to Japan, the U.S. is a low achiever, really. *Educational Researcher, 22*(3), 18–20.

Baker, D. P. (1997). Good News, Bad News, and International Comparisons: Comment on Bracey. *Educational Researcher, 26*(3), 16–17.

Barnett, M., & Finnemore, M. (1999). The Politics Power and Pathologies of International Organizations. *International Organization, 53,* 699–732.

Barro, S. M. (1996). *The International Education Expenditures Comparability Study* Washington D.C.: SMB Economic Research Incorporated (January).

Barro, S. M. (1998). *The Prospects for Developing Internationally Comparable Education Finance Statistics for Latin American Countries*. Washington, D.C.: SMB Economic Research Incorporated (March).

Beatty, A. (ed.) (1997). *Learning from TIMSS: Results of the Third International Mathematics and Science Study*. Washington, D.C.: National Academy Press.

Bennell, P. (1996). Using and Abusing rates of return: A critique of the World Bank's 1995 education sector review. *International Journal of Education Development, 16*, 235–248.

Biersteker, T. (1992). The Triumph of neoclassical economics in the developing world: policy convergence and the basis of governance in the international economic order. In J. Rosenau & E.O. Czempiel (Eds.), *Governance Without Government*. Cambridge: Cambridge University Press, 102–131.

Bloom, B. S. (1975). Implications of the IEA Studies for Curriculum and Instruction. In A. C. Purves & D. U. Levine (Eds.), *Educational Policy and International Assessment: Implications of the IEA Surveys of Achievement*. Berkeley: McCrutchan Publishing Corporation, 65–84.

Bloom, D. E. (2005). Global education: facts and data for measuring progress toward universal basic and secondary education. Paper prepared for the Project on Universal Basic and Secondary Education, Academy of Arts and Sciences, Cambridge, MA.

Bracey, G. W. (1996). International Comparisons and the Condition of American Education. *Educational Researcher, 25*(1), 5–11.

Bradburn, N., & Gilford, D. M. (Eds.) (1990). *A Framework and Principles for International Comparative Studies in Education*. Washington, D.C.: National Academy Press.

Bradburn, N., Haertel, E., Schwille, J., & Torney-Purta, J. (1991). A Rejoinder to 'I never promised you first place'. *Phi Delta Kappan* (June), 774–777.

Brown, G., & Micklewright, J. (2004). *Using International Surveys of Achievement and Literacy: A View from the Outside*, Montreal: UNESCO Institute of Statistics.

Brunner, J. J. (2005). Comparative Research and Public Policy: From Authoritarianism to Democracy. *Peabody Journal of Education, 80*(1), 100–107.

Carnoy, M. (1995). Structural Adjustment and the Changing Face of Education. *International Labour Review, 134,* 653–673.

Charbonnier, E. (2005a). Trends in the Financing of Education. Presentation to the Annual Meeting of the American Education Finance Association. Louisville, Kentucky, (March).

Charbonnier, E. (2005b). OECD Initiatives to Improve International Comparison of Education Statistics. Presentation to the Annual Meeting of the American Education Finance Association. Louisville, Kentucky (March).

Chromy, R.R. (2002). Sampling issues in design, conduct and interpretation of international comparative studies of school achievement. In National Research Council (Ed.), *Methodological advances in cross-national surveys of educational achievement*. Washington, D.C.: National Academy Press, 80–117.

Colclough, C. (1996). Education and the Market: Which parts of the neo-liberal solution are correct. *World Development, 24,* 589–610.

Colclough, C., & Mannor, J. (1991). *States of Markets: Neo Liberalism and the Development Policy Debate*. Oxford: Clarendon Press.

Coleman, J. S., Campbell, E. Q., Hobart, C. J., McPartland, J., Mood, A. M., Weinfall, F., & York, R. F. (1966). *Equality of Educational Opportunity*. Washington D.C.: U.S. Government Printing Office, Department of Health, Education and Welfare.

Corrales, J. (2006). Political Obstacles to Expanding and Improving Schooling in Developing Countries. In J. E. Cohen, D. E. Bloom, & M. B. Malin (Eds.), *Educating All Children: A Global Agenda*. Boston: MIT Press, 231–303.

Cummings, E., Knipe, H., Collins, J., Baker, D., & Gamson, D. (2007). The Increasing Rigor of Elementary Mathematics Textbook Content During the Twentieth Century. Paper presented to the Annual Meeting of the American Education Research Association, Chicago (April).

Dollar, D., & Pritchett, L. (1998). *Assessing Foreign Aid: What Works and What Doesn't and Why*. New York: Oxford University Press.

Elley, W. (1992). *How in the World Do Students Read?* Hamburg (Germany): Grindeldruck GMBH.

Floden, R.E. (2002). The Measurement of the opportunity to learn. In National Research Council (Ed.), *Methodological advances in cross-national surveys of educational achievement*. Washington, D.C.: National Academy Press, 231–267.

Gamoran, A., & Long, D. (2006). Equality of Educational Opportunity: A 40 Year Perspective. Wisconsin Center for Education Research Working Paper no. 2006–9, Madison, Wisconsin.

Gilford, D. M. (Ed.) (1993). *A Collaborative Agenda for Improving International Comparative Studies in Education*. Washington, D.C.: National Academy Press.

Glewwe, P., & Zhao, M. (2006). The Cost of Providing Universal Basic Education by 2015: An Evaluation of Cost Estimates. In J. E. Cohen, D. E. Bloom, & M. B. Malin, (Eds.), *Educating all Children: A Global Agenda.* Boston: MIT Press, 415–495.

Goldstein, H. (1995). *Interpreting international comparisons of student achievement*. Paris: UNESCO.

Gonzales, P., Guzman, J., Partelow, L., Pahlke, E., Jocelyn, L., Kastberg, D., & Williams, T. (2004). *Highlights From the Trends in International Mathematics and Science Study: TIMSS 2003* Boston.: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

Green, A., Preston, J., & Janmaat, J. (2006). *Education, Equality and Social Cohesion: A Comparative Analysis*. Hampshire: Palgrave Macmillan.

Guthrie, J. W., & Hansen, J. S. (eds.) (1995). *World Wide Education Statistics: Enhancing UNESCO's Role*. Washington, D.C.: National Academy Press.

Harris, D. (2007). Diminishing Marginal Returns and the Production of Education: An International Analysis. *Education Economics* (Forthcoming).

Helleiner, G. (1992). The IMF, The World Bank, and Africa's Adjustment and external debt problems: An unofficial view. *World Development, 20,* 779–792.

Heyneman, S. P. (1993a). Educational Quality and the Crisis of Education Research. *International Review of Education, 39*(6), 511–517.

Heyneman, S. P. (1993b). Comparative Education: Issues of Quantity, Quality and Source. *Comparative Education Review, 37* (November), 372–388.

Heyneman, S. P. (1995). The Economics of Education: Disappointments and Potential. *UNESCO Prospects XXV* (4) (December), 559–583.

Heyneman, S. P. (1997a). The Quality of Education in the Middle East and North Africa. *International Journal of Education Development, 17*(4), 449–466.

Heyneman, S. P. (1997b). Economic Development and the International Trade in Education Reform. *UNESCO Prospects, XXVII* (4) (December), 501–531.

Heyneman, S. P. (1998). Educational Cooperation Between Nations for the 21st Century, in J. Delors (Ed.) *Education for the Twenty-First Century: Issues and Prospects*. Paris: UNESCO.

Heyneman, S. P. (1999). The Sad Story of UNESCO's Education Statistics. *International Journal of Education Development, 19*(January), 65–74.

Heyneman, S. P. (2000). From the Party/State to Multi-Ethnic Democracy: Education and Social Cohesion in the Europe and Central Asia Region. *Educational Evaluation and Policy Analysis, 21*(4) (Summer), 173–191.

Heyneman, S. P. (2003a). What the United States Needs to Learn from UNESCO. *UNESCO Prospects, XXXIII* (1) (March),7–11.

Heyneman, S. P. (2003b). The History and Problems of making Education Policy at the World Bank 1960–2000. *International Journal of Education Development, 23*, 315–337.

Heyneman, S. P. (2004a). International Education Quality. *Economics of Education Review, 23,* 441–452.

Heyneman, S. P. (2004b). The Use of Cross-National Comparisons for Local Education Policy. *Curriculum Inquiry, 34*(3), 345–352.

Heyneman, S. P. (2004c). On the International Dimension of Education and Social Justice. *Journal of Education* (Boston University), *185*(3), 83–103.

Heyneman, S. P. (2005). Student Background and Student Achievement: What is the Right Question? *American Journal of Education, 112* (November), 1–9.

Heyneman, S. P. (2006a). The Effectiveness of Development Assistance in Education: An Organizational Analysis. *Journal of International Cooperation in Education* (Hiroshima), *9*(1), 7–25.

Heyneman, S. P. (2006b). The Use of Textbooks in a Modern Education System. In C. B. Bralavsky (Ed.), *Textbooks and Quality Learning for All: Some Lessons Learned from International Experiences*. Geneva: UNESCO?International Bureau of Education.

Heyneman, S. P. & Loxley, W. (1983). The Effect of Primary School Quality on Academic Achievement in Twenty-Nine High- and Low-Income Countries. *American Journal of Sociology, 88*(6) (May), 1162–1194.

Heyneman, S. P. & Fuller, B. (1989). Third World School Quality: Current Collapse, Future Potential. *Educational Researcher 18*(2) (March), 12–19.

Hinchcliffe, K. (1993). Neo-liberal prescriptions for education finance: unfortunately necessary or inherently desirable? *International Journal of Education Development, 13,* 183–187.

Horvitz, D. (1992). Improving the quality of international education surveys. Draft prepared for the Board on International Comparative Studies in Education. Washington, D.C.: National Research Council.

Ilon, L. (2002). Agent of Global Markets or Agent of the Poor? The World Bank's Education Sector Strategy Paper. *International Journal of Education Development, 22*(5), 475–483.

International Association for the Evaluation of Education Achievement (IEA). (1973). *Science Education in 19 Countries: An empirical Study*. Stockholm: IEA.

IEA. (1975). *Civic Education in Ten Countries, an Empirical Study*. Stockholm: IEA.

IEA. (1976). *The IEA Six Subject Survey: An Empirical Study of Education in 21 Countries*. Stockholm: IEA.

IEA. (1990). *Thirty Years of International Research: An Annotated Bibliography of IEA Publications (1960–1990).* Netherlands: IEA.

IEA. (1992a). *Computers in American Schools, 1992 An Overview*. Minneapolis: University of Minnesota, Department of Sociology.

IEA. (1992b). *How in the World Do Students Read? IEA Study of Reading Literacy*. Netherlands: IEA.

IEA. (1993). *Schools, Teachers, Students and Computers: A Cross National Perspective*. Netherlands: IEA.

IEA. (2001). *Citizenship and Education in 28 Countries: Civic Knowledge and Engagement at Age Fourteen*. Netherlands: IEA.

International Assessment of Educational Progress (1991). *The 1991 IAEP Assessment: Objectives for Mathematics, Science and Geography*. Princeton: Center for the Assessment of Educational Progress, Educational Testing Service.

International Labor Office (1996). *Impact of Structural Adjustment on the Employment and Training of Teachers*. Geneva: International Labor Office.

Jencks, C., Smith, M., Ackland, H., Bane, M. J., Cohen, D., Gintis, H., Heyns, B., & Michelson, S. (1972). *Inequality: A Reassessment of the Effect of Family and Schooling in America*. New York: Basic Books.

Linn, R. L. (2002). The Measurement of student achievement in international studies. In National Research Council (Ed.), *Methodological advances in cross-national surveys of educational achievement*. Washington D.C.: National Academy Press.

Martin, M., Gregory, K., & Stemler, S. (1999). TIMSS Technical Report. Boston International Study Center.

Medrich, E. A. & J. E. Griffith (1992). *International mathematics and science assessments: What have we learned*? Washington D.C.: U.S. Department of Education, Office of Educational Research and Improvement.

Merrett, S. (1971). Rate of Return to Education: A Critique. In R. Wykstra (Ed.), *Education and the Economics of Human Capital*. New York: Free Press, 196–213.

National Center for Education Statistics. (1996). *Education in States and Nations: Indicators Comparing U. S. States with Other Industrialized Countries in 1991*. Washington, D.C.: U.S. Department of Education.

National Commission on Excellence in Education. (1983). *A Nation at Risk: The imperative on education reform*. Washington, D.C.: U.S. Government Printing Office

National Research Council. (1995). International *Comparative Studies in Education: Descriptions of Selected Large-Scale Assessments and Case Studies*. Washington D.C.: National Academy Press.

National Research Council. (1999). Next Steps for TIMSS: Directions for secondary analysis. Washington, D.C.: National Academy Press.

National Research Council. (2002). *Understanding Others, Educating Ourselves: Getting More from International Comparative Studies in Education*. Washington, D.C.: National Academy Press.

Organization for Economic Cooperation and Development (OECD). (2002). *Education at a Glance 2002*. Paris: OECD.

OECD. (2004a). *Learning for Tomorrow's World: First Results from PISA 2003*. Paris: OECD.

OECD. (2004b). *Problem Solving for Tomorrow's World: First Measures of Cross-Curricular Competencies from PISA 2003*. Paris: OECD.

OECD. (2005). *Literacy, Economy and Society: Results of the First International Adult Literacy Survey*. Ottawa: OECD.

Olkin, I., & Searls, D.T. (1985). Statistical aspects of international assessments of science and education. Paper presented at the conference on Statistical Standards for International Assessments in Pre-College Science and Mathematics, Washington, D.C.

Peak, L. (1996). *Pursuing Excellence: A Study of Eighth-Grade Mathematics and Science Teaching, Learning, Curriculum and Achievement in International Context: Initial Findings for the Third International Mathematics and Science Study*. Washington, D.C.: National Center for Educational Statistics.

Peak, L. (1997). *Pursuing Excellence: A Study of U.S. Fourth-Grade Mathematics and Science Achievement in International Context: Initial Findings for the Third International Mathematics and Science Study*. Washington, D.C.: National Center for Education Statistics.

Phillips, D. (Ed.) (1991). *Lessons of Cross-National Comparisons of Education*. Oxford Studies in Comparative Education. Wallingford: Triangle Books Ltd.

Platt, W. J. (1975). Policy Making and International Studies in Educational Evaluation. In A. C. Purves & D. U. Levine (Eds.), *Educational Policy and International Assessment: Implications of the IEA Surveys of Achievement*. Berkeley, CA: McCrutchan Publishing Corporation, 33–59.

Postlewaite, N. (1975). The Surveys of the International Association for the Evaluation of Educational Achievement (IES). In A. C. Purves & D. U. Levine (Eds.), *Educational Policy and International*

Assessment: Implications of the IEA Surveys of Achievement. Berkeley, CA: McCrutchan Publishing Corporation, 1–33.

Postlewaite, N. (1999). *International Studies of academic achievement: Methodological Issues* Hong Kong: University of Hong Kong.

Pritchett, L. (2004). Access to Education. In A. M. Levine & M. J. Bane (Eds.), *Global Crisis, Global Solutions*. Cambridge: Cambridge University Press, 252–276.

Psacharopoulos, G. (1994). Returns to investment in education: a global update. *World Development, 22*(9), 1325–1343.

Psacharopoulos, G., Tan, J.P., & Jiminez, E. (1986). Financing Education in Developing Countries: An Exploration of Policy Options. Washington, D.C.: The World Bank.

Puryear, J.M. (1995). International Education Statistics and Research: Status and Problems. *International Journal of Education Development, 15*(1), 79–91.

Purves, A. C. & Levine, D. U. (Eds.) (1975). *Educational Policy and International Assessment: Implications of the IEA Surveys of Achievement*. Berkeley, CA: McCutchan Publishing Corporation.

Ravitch, D. (Ed.) (2003). *Brookings Papers on Education Policy.* Washington, D.C.: The Brookings Institution.

Rotberg, I. (1990). I Never Promised you First Place. *Phi Delta Kappan* (December), 296–303.

Rowen, B. (2002). Large-scale, cross-national surveys of educational achievement: Promises Pitfalls and possibilities. In A. C. Purves & D. U. Levine (Eds.), *Educational Policy and International Assessment: Implications of the IEA Surveys of Achievement.* Berkeley: McCrutchan Publishing Corporation, 321–353.

Schmidt, W., McKnight, C., & Raizen, S. (1997). *A Splintered Vision: An Investigation of U.S. Science and Mathematics Education.* Lansing, MI: National Research Center for the Third International Mathematics and Science Study, Michigan State University.

Sherman, J. D. (1981). Federal Education Policy in Australia: Implications for the U.S. School Finance project working paper. Washington, D.C.: National Institute of Education, (March).

Sherman, J. D. (1982). Public Finance of Private Schools: Observations from Abroad. Stanford Institute for Research on Educational Finance and Governance. Stanford, CA: Stanford University.

Smith, M.S. (2002). Drawing inferences for national policy from large-scale cross national education surveys. In National Research Council (Ed.), *Methodological advances in cross-national surveys of educational achievement*. Washington, D.C.: National Academy Press, 251–321.

Smyth, J. (1996). The Origins, Purposes and Scope of the International Standard Classification of Education. Paper submitted to the ISCED-revision task force. Paris: UNESCO (February) (Mimeographed).

Smyth, J. (2005). UNESCO's International Literacy Statistics 1950–2000. Paris: Mimeographed, 15.

Task Force on Higher Education and Society. (2000). *Higher Education In Developing Countries: Perils and Promise.* Washington, D.C.: The World Bank.

Task Force on Higher Education and Society. (2004). Moderated Discussion: The Task Force on Higher Education and Society. *Comparative Education Review, 48*(1), 70–88.

Thorndike, R. L. (1975). The Relation of School Achievement to Differences in the Backgrounds of Children. In A. C. Purves & D. U. Levine (Eds.), *Educational Policy and International Assessment: Implications of the IEA Surveys of Achievement.* Berkeley, CA: McCrutchan Publishing Corporation, 93–104.

Tomlinson, T., & Tuijnman, A. (eds.) (1994). *Education Research and Reform: An International Perspective*. Washington, D.C.: U. S. Department of Education.

UNICEF. (2002). *A League Table of Educational Disadvantage in Rich Nations.* Innocenti Report Card Issue 4 (November). Florence, Italy.

United Nations Education, Scientific and Cultural Organization (UNESCO). (1955). *World Survey of Education, Volume I: Educational Organization and Statistics.* Paris: UNESCO.

UNESCO. (1958). *World Survey of Education, Volume II: Primary Education.* Paris: UNESCO.

UNESCO. (1961). *World Survey of Education, Volume III: Secondary Education.* Paris: UNESCO.

UNESCO. (1966). *World Survey of Education, Volume IV: Higher Education.* Paris: UNESCO.

UNESCO. (1971). *World Survey of Education, Volume V: Educational Policy, Legislation and Administration Education.* Paris: UNESCO.

UNESCO. (2001a). *Latin American and the Caribbean Regional Report.* Paris: UNESCO.
UNESCO. (2001b). *Sub-Saharan Africa Regional Report.* Paris: UNESCO.
UNESCO. (2002). *Education for All: Global Monitoring Report 2002.* Paris: UNESCO.
UNESCO. (2004). *Education for All: Is the World on Track?* Paris: UNESCO.
UNESCO. (2005a). *Global Education Digest 2005. Comparing Education Statistics across the World.* Montreal: UNESCO Institute of Statistics.
UNESCO. (2005b). *EFA Global Monitoring Report 2005: The Quality Imperative.* Montreal: UNESCO.
Vinovskis, M. A. (1999). *The Road to Charlottesville: The 1989 Education Summit.* Washington, D.C. National Goals Panel.
World Bank. (1994). *Higher Education: The Lessons of Experience.* Washington, D.C.: The World Bank
World Bank. (1995). *Priorities and Strategies for Education.* Washington D.C.: The World Bank.
World Bank. (2005). *Achieving Education for All by 2015. Simulation Results for 47 Low-Income Countries.* Washington, D.C.: World Bank.