Achievement Inequality and the Institutional Structure of Educational Systems: A Comparative Perspective

Herman G. Van de Werfhorst and Jonathan J.B. Mijs

Department of Sociology and Anthropology, University of Amsterdam, Oudezijds Achterburgwal 185, 1012 DK Amsterdam, The Netherlands; email: H.G.vandeWerfhorst@uva.nl

Annu. Rev. Sociol. 2010. 36:407–28

First published online as a Review in Advance on April 20, 2010

The *Annual Review of Sociology* is online at soc.annualreviews.org

This article's doi: 10.1146/annurev.soc.012809.102538

0360-0572/10/0811-0407$20.00

Key Words

tracking, stratification, standardization, PISA, TIMSS

Abstract

We review the comparative literature on the impact of national-level educational institutions on inequality in student achievement. We focus on two types of institutions that characterize the educational system of a country: the system of school-type differentiation (between-school tracking) and the level of standardization (e.g., with regard to central examinations and school autonomy). Two types of inequality are examined: inequality in terms of dispersion of student test scores and inequality of opportunity by social background and race/ethnicity. We conclude from this literature, which mostly uses PISA, TIMSS, and/or PIRLS data, that inequalities are magnified by national-level tracking institutions and that standardization decreases inequality. Methodological issues are discussed, and possible avenues for further research are suggested.

INTRODUCTION

Thanks to the vast progress in data collection and availability, we are witnessing the rise of an elaborate, multidisciplinary literature on cross-national variation in student performance in primary and secondary schools. International data projects, such as the Program for International Student Assessment (PISA), the Progress in International Reading Literacy Study (PIRLS), and the Trends in International Mathematics and Science Study (TIMSS), have paved the way for a rich body of articles and official reports on international variation in average student performance, in the dispersion in performance, and in the influence of social origin and race/ethnicity on school performance.

An important research question that is addressed in this literature concerns the extent to which national educational institutions affect inequality in learning among students. Countries differ strongly in the organization of their educational systems, and it is important to know whether particular educational institutions are conducive to enlarging inequalities among students.

We focus on two types of inequality: inequality in learning as measured by the dispersion in test scores (which we label inequality as dispersion) and inequality of educational opportunity in terms of the influence of social class and race/ethnicity on students' test scores. These two aspects of inequality are conceptually distinct—an educational system could be relatively equal in terms of dispersion but unequal in terms of opportunities—yet, as we argue, theoretically and empirically linked.

With respect to educational institutions, we focus on the dimensions of differentiation in school types and tracks in secondary schooling systems and nationwide standardization. We see these characteristics of educational systems as institutions, as they reflect "laws, informal rules and conventions which give a durable structure to social interactions among members of a population" (Bowles 2004) in the form of (sub)national regulations on the way education is organized. Differentiation and standardization are both related to selection and allocation processes within and between schools.

We study differentiation primarily in terms of the external differentiation in separate school types and often schools. In the classification of types of curriculum differentiation of LeTendre et al. (2003), this review is mainly concerned with school-type differentiation and differentiation in terms of the geographical location of schools. Studies focusing on within-school-type ability grouping (i.e., internal differentiation) are not systematically reviewed (cf. Hopper 1968, LeTendre et al. 2003, Oakes 2005 [1985]). The reason for this omission is not that less institutionalized ability grouping would be irrelevant for inequality. Indeed, inequality in learning has often been observed in the American tracking system (Alexander et al. 1978; Ayalon & Gamoran 2000; Gamoran & Berends 1987; Gamoran & Mare 1989; Hallinan 1994a, 1996b; Lucas 1999, 2001), and tracking is prevalent in all educational systems in one form or another. Internal differentiation, however, is hard to capture in cross-national research, and few scholars have attempted to do so.

Standardization comes in different forms as well. First, the educational curriculum could be standardized across schools, with governments deciding what is to be taught in schools and which levels should have been achieved at which grades. The opposite side of standardization in this respect is school autonomy. Second, examinations can be standardized, as is the case in, for example, the Netherlands and England. Third, standardization can refer to the human and financial resources available to schools, e.g., in the form of teacher training and school budgets.

As some have noted, inequality (both in terms of dispersion and opportunities) is interesting particularly because of the potential trade-off between equality and efficiency of learning (Gamoran & Mare 1989, Hanushek & Wössmann 2005, Micklewright & Schnepf 2007, Wössmann 2008a). Educational systems that are well equipped to maximize average

student performance may be ill suited to guarantee equality of learning. We present these trade-offs in a broader perspective of what we see as the four core tasks of schooling. In the next section we discuss an institutional perspective of how national educational institutions may help or hinder the realization of those four core tasks. Then we discuss the data and research designs most commonly adopted by the empirical studies examined here, and address some methodological issues noted in the field. We also review the literature on inequality as dispersion and inequality of opportunity, before drawing conclusions and presenting some outlines for further research.

NATIONAL EDUCATIONAL INSTITUTIONS AND CORE TASKS OF SCHOOLING

To evaluate the relationship between educational institutions and inequality, it is important to position this relationship within a broader context of institutions and their effects. Therefore, we specify commonly agreed core tasks of education in contemporary Western societies. It is relevant to study those core tasks simultaneously to see whether a potential undesired effect of a particular educational institution (i.e., being detrimental for realizing a core task) has other desirable effects too.

We distinguish between four core tasks of schooling: (*a*) to offer/promote equality of opportunity; (*b*) to efficiently select and sort students on their abilities and interests; (*c*) to provide skills relevant for the labor market; and (*d*) to provide commitment to and skills relevant for active citizenship.

The first core task of schooling is to promote equal educational opportunities to children of different backgrounds (the equal opportunity task). Although we can develop separate measures for equality of opportunity and equality as dispersion, we should recognize three reasons why the two are linked. First, inequality of position can be linked to inequality of opportunities intergenerationally, as large inequalities in parental educational attainment lead to strong differences in their children's chances in school (Breen & Jonsson 2007). Second, one could maintain that equality of educational opportunity in terms of attained level is promoted if equality of position is granted at earlier stages in the educational career. Third, empirically it appears that equal societies in terms of dispersions are also more equal in terms of opportunities (Boudon 1974, Duru-Bellat & Suchaut 2005, Kenworthy 2008).

The second core task of schooling is to efficiently sort students according to their talents and interests (the efficiency task). Students are not equally equipped with different talents, and furthermore show interest in diverse kinds of work and living. Education can help to sort students in finding their career goals. The educational system should group students so that those with high learning skills have the opportunity to reach higher levels of schooling than those with lower learning skills. At the same time, those with other talents ought to have the chance to optimize their opportunities in the domains that more closely match their aptitudes. The total production of knowledge and skills is then deemed optimized (given a particular budget for education).

The emphasis on efficient learning by optimizing standardized test scores originates from an understanding of the role of education that has been developed largely from human capital theory. Human capital theory lends itself very well to education policy aimed at enhancing educational participation and improving academic achievement, as increasing inputs automatically lead to higher outputs (e.g., Hanushek 2006, Hanushek et al. 2008). Despite the evidence of a positive relationship between education and economic growth (but see Ramirez et al. 2006), the neoclassical model of education has been widely criticized (Bowles & Gintis 2000, 2002; Wolf 2002; Wolff 2006). However, given the noneconomic causal outcomes of schooling related to democratic citizenship, friendship networks, crime, and health (e.g., Becker 2007, Dee 2004, Hillygus 2005, Kubitschek & Hallinan 1998), even in the context of disputed causal effects

Table 1 Tasks of schooling and educational institutions

Educational institutions	Tasks of schooling			
	Promote equality of opportunity	Sort efficiently	Prepare for employment	Prepare for active citizenship
Differentiation				
Standardization				

of schooling on economic returns policy makers may still want to maximize learning in schools.

A third core task is that education should prepare students for the labor market (the labor market task). This task implies that education provides skills that are productive for work, thereby aiding graduates in optimizing their labor market opportunities and employers in optimizing production. One way to promote the labor market task of schooling is by promoting efficiency: The level of skill is optimized, which thereby benefits the economy. Yet labor market preparation may imply the optimization not only of the average level of skill, but also of the particular distribution of a type of skill, i.e., vocationally specific or general. Educational systems differ widely in the emphasis placed on vocational or general skills. The German educational system can, for example, be seen as strongly vocationally specific, whereas the American and Scandinavian systems can be seen as more vigorously emphasizing general skills.

The fourth central task of educational institutions is to stimulate citizens to become actively involved in society (the active citizenship task). Educational institutions, it follows, should be seen as more than sites of production of work skills. Their output should include the formation of critical citizens who can actively take part in social and political life (Ten Dam & Volman 2003, 2007; Terwel 2005; Torney-Purta et al. 2001). This policy target implies that the formation of active citizenship cannot be delegated fully to private caregivers. Schooling could have an active role in it, as it can help promote equality in terms of civic competence and commitment. Civic skills overlap partially with the skills demanded by the labor market. For instance, general cognitive skills contribute to labor market productivity as well as to being informed about politics and, thereby, to political participation. Partly, though, civic skills and commitment are created in subjects specifically devoted to knowledge of politics, of formal institutions, and of current affairs, subjects that may not contribute to learning skills relevant for work.

We may think of ways in which different kinds of educational institutions affect the likelihood that the four core tasks of education can be achieved. In this review, we focus on two educational institutional variables: the extent of differentiation in school types and the level of national standardization.[1] For heuristic reasons we plot these two dimensions of educational systems against the four central tasks of schooling systems. This heuristic framework allows us to examine which relationships exist between educational institutions and the tasks of schooling (see **Table 1**). When, within a row, a particular educational institution affects the achievement of one particular central task of education positively, and at the same time the achievement of another central task negatively, we have detected a policy trade-off.

If we inspect the first two of these tasks, equality and efficiency, we encounter a central policy trade-off in educational design. In systems where students are selected and differentiated into separate schools as early as age 10 (Germany) or age 12 (the Netherlands), inequality may be higher than in systems

[1]One could add that countries vary in the extent to which students have the opportunity to move between tracks (Gamoran 1992, Hallinan 1996b, Kerckhoff 2001), but few scholars have examined this variability cross-nationally.

without external differentiation (e.g., Sweden or the United States), both with regard to dispersion of achievement scores and social-structural effects on achievement. Yet, although inequality may be higher in differentiated systems, efficiency of learning may also be higher. When all students benefit from being placed in homogeneous classes, or when high-performing students benefit more than the low-ability students lose, the average performance may be higher in strongly stratified educational systems. Some of the studies reviewed here have explicitly examined this trade-off empirically.

The same equality-efficiency trade-off can also be detected with regard to the standardization of an educational system. Standardized educational systems can be helpful in reducing inequalities by implementing central examinations that temper parental influence. Yet a standardized system with regard to national curricula may impede competition between schools, which could lead to lower achievements of students.

Another trade-off is the one between labor market preparation and equality of opportunity. A stratified educational system (in particular, one with a strong vocational sector) clearly helps youngsters in the transition process from the educational system to work (Arum & Shavit 1995, Breen 2005, Müller & Gangl 2003, Shavit & Müller 1998). Yet there is a significant social class effect on choice of school type. If students enrolled in vocational tracks have fewer opportunities to access tertiary education, strongly vocationally oriented systems may enlarge social class differences in the attainment of a tertiary-level degree (OECD 2008).

A less well-known trade-off is that between labor market preparation and citizenship education. It may be that educational systems that perform well in the preparation for the labor market perform worse when it comes to the citizenship task of schooling. This could be the case if early selection in the educational system, and a strong vocational orientation, both lead to improved labor market signaling but at the same time increase variation in citizenship competences. A well-known view on social justice is that principles of justice depend on the domain of life we are looking at (Miller 1999, Walzer 1983). If we subscribe to the view that we should abide by an equality principle in issues related to politics and that a merit principle should govern the labor market (cf. Miller 1999), we then easily see that educational systems are functioning precisely at the crossroads of these principles. Schooling, that is, should aim at both increasing the visibility of the merits/skills obtained, which is promoted by differentiation, and at equalizing the kind of knowledge and skills that promote democratic equality, on which differentiation could have detrimental effects.

DATA AND METHODOLOGY

Data

A vast and growing number of studies take an internationally comparative view to the study of educational inequality. In this section we briefly describe the three most popularly studied data sets: the PIRLS, TIMSS, and PISA data. All of these data sets are collected in a wide range of countries, both Western and non-Western.[2]

The Progress in International Reading Literacy Study (PIRLS) concerns students in fourth grade of primary school, i.e., children of nine or ten years of age. Students are tested on reading skills and are asked to answer questions on their situation at home and at school. Also, their parents and teachers are asked to provide information concerning the child's development, their own role in upbringing and educating, and the school environment. Finally, the school board provides information on the schools' organization and staff and on the characteristics of the student population. This information provides insight into differences in test scores as well as potentially influential factors

[2]For further information on the countries covered, we refer to the Web site of the International Association for the Evaluation of Educational Achievement (IEA) which runs the TIMSS and PIRLS data sets (**http://www.iea.nl**), and the Organisation for Economic Co-operation and Development (OECD) that collects the PISA data (**http://www.pisa.oecd.org**).

such as school environment, student characteristics, and family background.

To enable cross-country comparison, test scores are standardized to an international average of 500 with a standard deviation of 100. The data are collected through a two-stage stratified sampling design: Schools are selected in the first round of sampling, and classes within those schools are selected in the second round. The target is to select a minimum of 150 schools for all participating countries, of which one class per school is sampled.

The first PIRLS was conducted in 2001 and a second in 2006, and the intention is to maintain this five-year cycle. Thirty-five countries participate in PIRLS, with 150,000 respondents in 2001. The test is financed jointly by the participating countries and the World Bank and is performed under the supervision of the International Association for the Evaluation of Education Achievement (IEA), an organization of 62 countries that has been active since 1959. The same organization also coordinates the TIMSS.

The Trends in International Mathematics and Science Study (TIMSS) tests children in grade seven or eight. The questions, on mathematics and science, are based heavily on the standards of international curricula—the common ground of the participating countries' curricula. The test scores (average 500, standard deviation 100) are supplemented by background questions for students and questions on factors specific to schools, which are addressed by the school board. The sample is drawn in like manner as that of PIRLS: At least 150 schools per country are selected, of which one grade seven and one grade eight class is randomly picked. The first TIMSS was conducted in 1995, and the test was repeated every four years, in 1999, 2003, and 2007.

The Program for International Student Assessment (PISA) tests students 15 years of age on their knowledge and skills in reading, mathematics, and science. Scores are standardized in like manner as that of PIRLS, albeit with differences in what is measured: Whereas PIRLS tests only reading and TIMSS is based on an international standard (curriculum), PISA is aimed toward assessing general skills and competencies related to real-world situations.

The data collection of PISA, through a two-stage stratified sampling, is organized somewhat differently from that of PIRLS and TIMSS, as the target group is an age group rather than a grade level. This choice may have implications for inequality as countries differ in the extent to which students repeat classes in case of underperformance. In countries where grade repetition is more common, a relatively larger share of students of one age cohort is educated in lower grades than in countries where repetition is rare (Raudenbush & Kim 2002). This may affect their performance negatively, potentially leading to stronger social class effects if repetition is unevenly distributed across social classes.

After schools have been randomly selected in the first stage, individual students are selected in the second. PISA is an initiative of the Organisation for Economic Co-operation and Development (OECD) and has been repeated every three years since 2000. In that year 43 countries participated; 41 did so in 2003, 57 in 2006, and 69 in 2009. The number of respondents is 250,000 and above.

Research Designs

The literature discussed in this review can be divided roughly according to two methodological approaches: studies that analyze cross-sectional data, linking educational inequality in a country to specific characteristics of that country's educational system, and studies adopting a difference-in-difference (DiD) model. The latter studies have been designed to deal with the typical endogeneity problem of cross-sectional surveys; the enrollment in a track is endogenous on school performance (and social class). DiD models compare inequality measured in two data sets at different points in time—commonly, one data set based on primary education (i.e., before selection) and one on secondary education (i.e., after selection in tracks). If differentiation has a negative effect on

equality, inequality is assumed to increase more across educational transitions in those countries that have a large degree of differentiation. The first data set (i.e., primary school) serves as a point of reference for the analysis of the second data set, effectively holding constant any unobserved differences between countries.

Although the above-mentioned property of DiD models surely is one that appeals, the approach suffers from one important weakness: comparing different data sets assumes that educational ability or performance is measured in like manner, whereas in fact it is not (Micklewright & Schnepf 2007). Ammermüller (2005) formalizes this point by stating that the measure derived from the second data set contains an error component when used as a second measure of the first data set. Ammermüller's formalization, however, makes it possible to formulate assumptions under which a DiD model is worth using. The most important of these assumptions, in research on inequality of educational opportunity, is that there is no variation between countries in the correlation between the error component and unmeasured aspects of family background. This may not be a very plausible assumption, as countries are likely to vary systematically in the way that children are sorted on the basis of noncognitive characteristics. For instance, in unstandardized educational systems it is plausible that social background matters more for track placement relative to standardized systems, precisely for reasons other than cognitive ones.

There are other techniques for modeling the unobserved variation between countries in cross-sectional data, by using a dummy variable for countries in mixed models (Brunello & Checchi 2007). This leads us to regard both approaches as important sources for our review.

Methodological Issues

Several methodological issues have been raised (see Porter & Gamoran 2002b for a comprehensive overview). We discuss three methodological issues that warrant particular attention for our review on institutions and inequality.

First, there has been some debate as to the extent to which standardized tests can be used to compare educational systems. The OECD claims that PISA tests knowledge and skills for life and is not meant to measure whether students have mastered a particular curriculum (OECD 2004, Kirsch et al. 2002). Yet, as Goldstein (2004) notes, PISA claims to evaluate the performance of educational systems, which seems at odds with the curriculum-independent focus of PISA.

Moreover, an evaluation of educational systems is difficult if one depends on cross-sectional data like PISA, TIMSS, and PIRLS. In national educational studies it is commonplace to use longitudinal data to examine the contribution of schools to learning. One needs information on earlier demonstrated ability to draw conclusions with regard to the causal impact of schooling on learning.[3] Also in a comparative framework one would prefer longitudinal data to examine whether educational systems affect (inequality in) learning. Inequality of educational opportunity results from two sorts of inequality: inequality in achievement (often called primary effects; Boudon 1974, Erikson et al. 2005, Goldthorpe 1996) and inequality in educational choices conditional on achievement (called secondary effects). Separating primary from secondary effects requires longitudinal data, where educational choice at a certain point in the school career can be modeled to be conditional on achievement at an earlier point in time.[4] Especially if one is interested in cross-national variations in inequality, one would be particularly interested to see whether particular educational institutional characteristics affect mostly primary or secondary effects.

[3] Note that, even if longitudinal data are available, an achievement growth may be attributed to schooling but could also be attributable to other (e.g., family) influences.

[4] The separation of primary and secondary effects is, however, also somewhat problematic, as a student's achievement may depend on the choices he or she is expecting to make in the future.

This has not been possible in the comparative surveys to date.

Second, the DiD models used to simulate a longitudinal research design are not without problems. The analysis of country-level data prohibits the disentanglement of a country's variance into between-school and within-school components. Additionally, a comparison of two surveys assumes that the tests are comparable, which rests on the assumption of one underlying distribution of, say, mathematical proficiency (assuming no measurement error). Yet precisely this feature of achievement tests has been disputed.

That brings us to the third methodological issue—the validity of the assumptions underlying the psychometric tests implemented in the surveys. Are these assumptions plausible given the findings on cross-national variation in average scores and dispersion in test results? The standardized tests in the various surveys have been developed on the basis of Item Response Theory (IRT). There are two assumptions in IRT that are relevant for our study of institutional effects. First, IRT assumes that there is an underlying true distribution of a student's proficiency in a particular subject (OECD 2005, Yamamoto & Kulick 2000). Second, the underlying true distribution has no fixed range but could, theoretically, reach infinity. The distribution is assessed by repeated measurement, offering various tests consisting of different items with different levels of difficulty. Under the assumption that there is a true distribution of proficiency (or ability) independent of the test, it is possible that extremely bright students have a proficiency level above the threshold of the most difficult item. These students may be "assumed to be capable of at least the achievements described for the highest level" (OECD 2005, p. 256). Furthermore, if the proficiency level of students is below the lowest level defined, a student's proficiency is "lower than that which PISA can reliably assess" (OECD 2005). These quotes illustrate that the PISA tests are seen as measuring an underlying distribution that exists independent of the test and that the true proficiency level of a student may substantially exceed the highest defined level, illustrating an unbounded upper tail of the distribution.

Several authors have criticized the assumption that such a true distribution of achievement or ability exists independently of the test. Atkinson (1975, p. 89; see also Mayer 1960 and Micklewright & Schnepf 2007, p. 133) noted that "the distribution [of ability] depends on the measurement rod used and cannot be defined independently of it." Given that each survey uses its own test, strictly speaking each survey "aims to measure somewhat different things" (Brown & Micklewright 2004, p. 42). Brown et al. (2006) examined the country medians and dispersions and found that country-level correlations between different domains were higher within surveys (e.g., reading and mathematics within PISA) than were correlations within domains between surveys (e.g., mathematics in PISA and mathematics in TIMSS). This finding made the authors cautious about interpreting the different surveys as measuring one underlying concept of proficiency in reading, mathematics, or science. Another study, however, argues that the country-level correlations on the different tests are so large that one could speak of a general intelligence (*g*) factor underlying all tests (Rindermann 2007).

The assumption of infinity in the underlying proficiency has not been discussed in the literature. We do not think this assumption is too important; for the understanding of averages and inequalities in achievement, a test with an upper (and lower) bound is also valuable. However, the empirical results on the nonexistence of a trade-off between equality and efficiency may be related to the fact that there is an upper limit in achievement. Similar to educational attainment, which more evidently has an upper limit, higher averages seem to coincide with lower dispersions (e.g., Hauser & Featherman 1976, Rijken 1999, Thomas et al. 2001). Decreasing dispersions with increasing average achievement test scores may possibly be due to the fact that there is an upper limit in the distribution of achievement, similar to attainment.

Despite these methodological concerns, the data are of high quality. As Porter & Gamoran (2002b, p. 8) conclude in the volume that they edited about cross-national student achievement data, "the level of methodological quality is high, and therefore the findings of large-scale studies are worth taking seriously."

EMPIRICAL FINDINGS I: INEQUALITY AS DISPERSION

To what extent do countries vary in the dispersion in achievement tests, and how/to what extent is the level of dispersion related to educational institutional characteristics?

Hanushek & Wössmann (2005) examine inequality as dispersion through a DiD approach. The authors analyzed learning outcomes in primary education (grade four) using PIRLS and TIMSS data and those in secondary education (grade eight and age 15) using TIMSS and PISA data, respectively. They found, for each measure of dispersion used, that once the variance in primary school is controlled for, the variance in learning increased more in those countries in which students make the transition from primary school to a highly differentiated secondary school, relative to students who stay in a comprehensive system. In other words, controlling for the fact that countries, for unknown reasons, have different levels of variance in achievement in primary school, early selection leads to an increase in inequality at the secondary school level.

Huang (2009) approaches the matter in a different way. His study takes as the independent variable classroom homogeneity, measured as the between-school component of the total variance in students' mathematics scores. This, he argues, is at the core of the theoretical explanation of why differentiation contributes to efficiency of learning. Teachers, the explanation holds, can teach more effectively when their students are of comparable ability levels (cf. Hallinan 1994b). In a study of TIMSS data for 24 countries, Huang shows that this explanation does not hold. The effects of classroom homogeneity on inequality as dispersion are strong and indicate that high-performing students stand to gain at the expense of low-performing students.

However, findings of other studies are not in line with the hypothesis of increasing variations among students with increasing differentiation. Micklewright & Schnepf (2007) study inequality as dispersion using the cross-sectional TIMSS, PISA, and PIRLS data sets. Dispersion is measured by differences between achievement scores of students in different parts of the distribution (5th, 50th, and 95th percentiles). Micklewright & Schnepf show that even in a low-dispersion country like France, the difference between P95 and P5 is 5.5 times as large as the average gain of one year of schooling. In countries with higher dispersions, this factor goes up to 10.7 in Germany and 13.3 in the United States. It may thus be concluded that dispersion in learning is substantial and that variations among countries are to be taken seriously.

Although Micklewright & Schnepf (2007) find strong cross-national variation between countries in terms of dispersion, our own analysis of their findings shows no positive relationship between differentiation institutions and dispersion. We calculated the correlation between a country's level of dispersion and various indicators of differentiation: the number of tracks available to a typical 14-year-old student, the length of the tracked curriculum as a proportion of the total length of secondary education, and the percentage of students enrolled in vocational education at upper secondary and lower secondary levels (taken from the OECD's Education at a Glance reports, and Brunello & Checchi 2007). These correlations were all around −0.20, indicating lower dispersions when differentiation in the system increases. However, none of these correlations turned out to be statistically significant (with *N*s between 15 and 18 countries). Also Duru-Bellat & Suchaut (2005) find no significant statistical relationship between tracking at age 14 and dispersion in test scores, which they examined by looking at the proportion of students classified in the top and

bottom categories of the standardized PISA ranking.

Another study that refutes the hypothesis that external differentiation leads to larger dispersions in learning was done by Vandenberge (2006). He examines dispersion in the residuals of a regression equation predicting performance, controlling for social background. However, unlike other studies, Vandenberge uses enrollment in the vocational track at age 15 as an indicator of differentiation. The latter indicator is unfortunate because placement in a vocational track at age 15 is unlikely to be a good representation of the level of differentiation of the system. In Vandenberge's study, Germany, for example, scores extremely low on the indicator of differentiation because German vocational apprenticeships are offered after the age of 15. Given the strongly selective nature of the German educational system, with selection around the age of 10 into separate schools, it may be more sensible to measure differentiation by age of first selection, regardless of vocational enrollment. Indeed, Vandenberge finds more evidence for an impact of differentiation with regard to the effect of interschool segregation, which is measured by the standard deviation in performance across schools within countries. However, as a concept, interschool segregation confounds segregation from a source independent of the educational structure (i.e., neighborhood segregation) with segregation imposed on the system through the structure of selection and school organization. Thus, we are left unsettled with regard to the effect of educational institutions. In summary, when it comes to the relationship between external educational differentiation and inequality as dispersion, the evidence is mixed.

Some of the studies mentioned above have also analyzed the existence of a trade-off between equality and efficiency. It appears that more dispersion goes hand in hand with a lower median score. Thus, no support is found for the hypothesized trade-off between the mean and variance of learning achievement. It is not the case that more efficient learning can be obtained by institutions that allow for greater inequality among students (Brown et al. 2006, Hanushek & Wössmann 2005, Micklewright & Schnepf 2007). Furthermore, differentiation may reduce efficiency in another way because it poses a serious barrier for entry to higher education (Van Elk et al. 2009, Marginson et al. 2007). Enrollment in tertiary education tends to be substantially lower in countries with strongly externally tracked systems.

In a study of seven Central and Eastern European (CEE) countries, Ammermüller et al. (2005) find that all four of the well-performing CEE countries (Hungary, Czech Republic, Slovakia, and Slovenia) have tracked educational systems and higher levels of inequality, whereas the three worst-performing countries (Lithuania, Latvia, and Romania) all have comprehensive educational systems and lower levels of inequality. Thus, among CEE countries, there is evidence of a trade-off between efficiency and equality in school performance.

Besides these cross-national studies, some scholars have studied institutional change within countries from a differentiated to a comprehensive system. Their research shows, for a number of countries, that a development toward more comprehensive schooling increases the average performance and reduces inequalities (Duru-Bellat & Kieffer 2000, Gamoran 1996, Gamoran & Weinstein 1998, Meghir & Palme 2005, Pekkarinen et al. 2009).

If we think of the mechanisms that explain why educational differentiation increases variability between students but does not increase the average performance substantially, we may learn from the single-country studies on school compositional effects. Although effects of school (or class) composition exist independent of the system of differentiation, tracking institutions affect the within- and between-school variation in school achievement and social class. In strongly differentiated systems, variability between schools is relatively high, and within-school variability relatively low (OECD 2004).

A few findings are relevant to explain the cross-national findings discussed above. First, tracking affects average performance more for high-ability (high-track) students than for

low-ability (low-track) students, for whom the effects of tracking are often negative (Hallinan 1988, Hallinan & Kubitschek 1999, Hattie 2002, Huang 2009, Kulik & Kulik 1982, Marsh 1984). Also, low-ability students benefit more from mixed-ability grouping (Dobbelsteen et al. 2002, Thrupp et al. 2002, Van den Eeden & Terwel 1994, Willms 1986, Zimmer & Toma 2000). Second, even for high-ability students the gains of tracking are not so great, often around 0.10–0.15 standard deviations in experimental research designs. Third, the overall effect of tracking on student performance is very small and often negative (Figlio & Page 2002, Hattie 2002, Michaelowa & Bourdon 2006, Thrupp et al. 2002). There is thus little evidence that tracking positively affects efficient learning, although it increases variability in achievement between students of different levels of demonstrated ability (Betts & Shkolnik 2000, Hattie 2002, Huang 2009, Zimmer & Toma 2000). One explanation suggested for the limited gains from tracking is that teachers do not adjust their methods of instruction when they teach homogeneous classes, and when teachers do, it may particularly slow down the instruction to low-ability students in a way that impedes their future learning (Gamoran & Berends 1987, Hattie 2002).

EMPIRICAL FINDINGS II: INEQUALITY OF OPPORTUNITY

Differentiation and Family Background Effects

Since the 1960s, educational systems have been typified with regard to their impact on inequalities (Hopper 1968, Turner 1960). Some cross-national empirical studies on achievement focused on institutional and structural factors other than educational that affected inequalities in learning (e.g., Baker et al. 2002, Heyneman & Loxley 1983). Other early studies have, however, focused on early selection in comparison to comprehensive schooling, pointing to larger effects of social background in systems where students are selected at an earlier age (Comber & Keeves 1973; Husen 1967, 1973). More recent studies on the impact of educational differentiation on equality of opportunity by family background confirm these earlier findings: Inequalities are magnified by external educational differentiation. Following is an in-depth discussion of these recent studies. Fortunately, all of the cross-national student achievement surveys collect information on family background, although more homogenization of measurement instruments has been called for (Buchmann 2002).

Schütz et al. (2008) focus on differentiation as a potential explanation for cross-national differences in inequality of opportunity. The number of books in the household, which the authors take to be the best reflection of family income, has a stronger effect on student performance in countries that track students in different school types on the basis of ability. Note that sociologists tend to argue that cultural possessions such as books are a reflection more of cultural than of financial resources available in the household, and cultural climate tends to have a stronger effect on student achievement than financial resources have (Marks 2005, Park 2008a). Studying institutional effects more comprehensively than other studies have done, Marks (2005) shows that the explanatory power of social class on literacy is higher in countries with lower participation rates in university education, more educational tracks at age 15, an earlier age of first selection in the schooling system, a higher between-school variance in learning outcomes, and more income inequality (although the latter relationship is modest). Thus, there is clear evidence that differentiation, measured in various ways, increases inequalities in learning by social class. Horn (2009) comes to the same conclusion in his cross-national study of 29 countries using PISA 2003 data, and Bauer & Riphahn (2006) confirm these findings as well for the 26 cantons of Switzerland that vary in timing of tracking.

Wössmann (2008a) makes use of TIMSS 1995 data to assess the effect of family background (parental education and number of books in the household) on students'

mathematics achievement. His study shows that in the United States and in all 17 European countries included in his study, family background is correlated with a student's performance in mathematics. The effect varies from relatively weak in Portugal, France, and Wallonia (Belgium), to strong in England, Germany, and Greece. When comparing the mathematics scores of students from highly educated parents to those with only primary education, the differences are highest in the United States—higher than in any European country.

Importantly, Wössmann (2008a) and Schütz et al. (2008) show that there is no relationship between social inequality of learning and average school performance, which refutes the hypothesis of a possible trade-off between efficiency and equality. Elsewhere Wössmann (2008b) argued that a trade-off between efficiency and equality of opportunity, if at all likely, is to be found only at the higher stages of education. From the life-cycle model of human capital formation, which shows that early learning enhances later learning (e.g., Heckman 2000), it follows that it would be more efficient to offer further education to those with high levels of early performance. This would enhance inequality in learning. At the early stage of (pre)schooling, however, more gains can be expected for children from disadvantaged backgrounds, which implicates a complementarity, rather than a trade-off, between efficiency and equity.

Ammermüller (2005) has conducted a DiD approach to study inequality of educational opportunity for 14 countries using PIRLS (2001) and PISA (2000) data. For all family background variables, this study finds that the effects increase most strongly in those countries that have either a highly differentiated educational system or a large private school sector. The yearly amount of instruction at school reduces the effect of family background, whereas the level of autonomy for schools increases the effect of parents' educational stance on their children's reading skills.

Brunello & Checchi's (2007) research is the most comprehensive study on the relationship between differentiation and inequality of opportunity by family background. The authors used multiple data sets, on students as well as on (young) adults, to analyze the effect of family background on educational and on labor market opportunity. Differentiation is operationalized in two ways: length of school-type differentiation within the educational system and percentage of students enrolled in secondary vocational education. In addition to the standardized test scores used in most research, Brunello & Checchi include indicators of educational attainment, access to tertiary education, school dropout rates, language and mathematical skills, as well as indicators of early labor market experience such as employment, training, and income. Their data sources are the European Community Household Panel (ECHP), the International Social Survey Project (ISSP), the International Adult Literacy Survey (IALS), and PISA 2003.

The authors show how, for a number of outcomes, the effect of family background increases with length of tracking. This concerns the effect of family background on educational attainment, school dropout rates, access to and enrollment in tertiary education, and job income. With regard to mathematics, however, the outcomes for young adults using IALS data contradict the hypothesis that differentiation increases inequality of opportunity; family background's positive effect on test scores decreases as tracking length increases. The authors offer two possible explanations for the discrepancy between their findings for mid-teen students and young adults. First, time spent in the labor market could compensate for the negative effect of tracking on equality of opportunity. IALS respondents often have some labor market experience, and those in countries with early tracking tend to have more experience as they are less likely to enroll in tertiary education. A second explanation is that the full effect of tracking can better be assessed at a later age (IALS) than at age 13 or 15.

With regard to lifelong learning, the authors find also that the positive effect of family background decreases as differentiation increases. In

other words, although students from highly educated parents are more likely to receive a form of training anywhere, this effect is stronger in countries with a comprehensive system of education than in those with highly differentiated educational systems. Importantly, Brunello & Checchi conclude that the length of school-type differentiation increases inequality of opportunity but that the vocational orientation does not.

Differentiation and Race/Ethnicity Effects

In addition to the effects of family background, race/ethnicity effects are sometimes found to depend on the educational institutional structure. School tracking has been related to ethnic inequality in European cities and countries (Crul & Holdaway 2009; Crul & Schneider 2007, 2009; Crul & Vermeulen 2003). Israeli studies have shown that tracking impedes the educational opportunities of Arab and Jewish minorities relative to the majority Jewish ethnic group (Shavit 1984, 1990). Entorf & Lauk (2008), furthermore, showed that differentiated educational systems tend to magnify previously existing inequalities in educational performance between migrant students of low socioeconomic background and students from more privileged families. Peer effects in high-ability tracks favor students with an initial advantage, whereas disadvantaged students experience negative peer effects in the lower tracks (a finding confirmed by ethnographic accounts; e.g., Paulle 2005). Immigrant students, the authors report, would benefit from a more diverse student population in terms of educational ability within, rather than between, schools.

Race/ethnicity has also been associated with track mobility, although only from a single-country perspective. Hallinan (1996a) finds that students of color in the United States, in general, are disproportionately assigned to lower tracks in both English and mathematics and are more likely than white students to drop out of these tracks. Studying track mobility in the Netherlands, Kalmijn & Kraaykamp (2003) constructed an event-history model for secondary school careers and showed that students of ethnic minorities were less likely than nonminority students to experience downward track mobility, although they were more likely to drop out of school before completion.

In the North American context, race is correlated strongly with indicators of student achievement. Studies show that African American students are disproportionately assigned to lower tracks (Darling-Hammond 1994; Hallinan 1991, 1992; Lucas 1999; Oakes 1990). When student achievement is controlled for, however, the racial effect is ambiguous. Some studies find that race differences in track assignment diminish or disappear (e.g., Hallinan 1994a, Pallas et al. 1994), whereas others indicate that minorities are favored in the assignment to higher-ability groups (Gamoran & Mare 1989, Stanat & Christensen 2006, Van de Werfhorst & Van Tubergen 2007). An unsettled matter, despite receiving particular attention in the literature, is whether educational tracking has negative ethnicity effects on inequality of educational opportunity independent of social class effects. In addition to the previously cited studies, Alexander & Cook (1982) find no evidence for such an effect, although Hallinan (1994b) in a later study does.

Standardization and Family Background Effects

It is important to examine the repercussions of standardization alongside differentiation, as standardization makes the process of grouping students transparent and objectifies its criteria. Indicators of standardization, such as central examinations, a national curriculum, and standardized school resources (regarding budgets, staffing decisions, teacher training, establishing salaries), reduce the influence of social origin on student performance (although more so in TIMSS than in PISA), whereas school autonomy increases the influence of social origin (Horn 2009; Muller & Schiller 2000; Park 2008b; Schütz et al. 2008; Wössmann 2003a,b, 2005). Additionally, standardized

educational systems are said to reflect that qualifications represent the same skill level throughout the country, implying stronger signaling to the labor market (Shavit & Müller 1998). However, standardization is not significantly related to intergenerational mobility in educational attainment, as has been shown with IALS data (Pfeffer 2008).

An interesting study to examine in more detail is Stevenson & Baker's (1991, p. 1) investigation of what they call the implemented curriculum—"that portion of the curriculum that is taught to students in the classroom"—with an interest in explaining variability by addressing the level of standardization (state control over the curriculum). In their analysis of SIMS data (Second International Mathematics Study, the predecessor of TIMSS) for 15 educational systems, they find that standardization decreases the effect of student and teacher characteristics on how much mathematics is taught in class. Students in standardized school systems, that is, are more likely to get the same education, whereas school systems with more school, local, or provincial autonomy exhibit higher variability within and across schools in what students are taught. Their study is criticized by Westbury & Hsu (1996), who argue that it is not standardization but within-country differences in school types (external differentiation) or within-school tracks (internal differentiation) that drive the results.

With regard to efficiency, school autonomy boosts average performance only in countries with central examinations (Dronkers & Robert 2008, Fuchs & Wössmann 2007). So, in the context of central examinations warranting school accountability, there is a potential trade-off between equality and efficiency, for autonomy tends not only to enhance average performance, but also to increase the levels of inequality of educational opportunity. However, in the context of more limited levels of external accountability, we may tentatively conclude that there is no such trade-off, as the efficiency gains of school autonomy are much less pronounced.

Studying educational differentiation and standardization jointly is relevant for two reasons. First, as both indicators may be related, one needs to control for standardization to assess the impact of differentiation, and vice versa. Pfeffer (2008) examines both institutional characteristics with regard to educational mobility between parents and children with regard to their highest attained level of schooling. His results show that differentiation hinders educational mobility but that standardization has no effect at all. Also, Horn (2009) analyzes the impact of differentiation and standardization simultaneously using PISA data and shows that inequality of opportunity increases with differentiation and school autonomy and that uncontrolled effects are largely the same as controlled effects due to the low correlation between standardization and differentiation indicators.

Second, it has been argued that there is an interaction effect between differentiation and standardization. In a comparison between Israel and the United States, Ayalon & Gamoran (2000) demonstrate that similar educational reforms (i.e., differentiation) can lead to different outcomes due to differences in the degree of the country's level of standardization. Israel's standardized curriculum and national examinations appear to be an incentive for achievement among teachers and students in all levels of schooling, whereas their absence in the United States reinforces inequality without raising average scores. In other words, in a standardized school setting with limited autonomy of schools with regard to standards, curriculum development, and teacher training, the effect of tracking on inequalities may be tempered relative to decentralized systems. Such an effect is also suggested by Broaded (1997) in his study of Taiwan, where he found very limited influence of family background on educational aspirations and senior high school placement. His results indicate that the standardized curriculum of the Taiwanese educational system, as well as its practice of standardized examination, provide the institutional context in which tracking helps realize the efficiency target without negatively affecting inequality of

Table 2 Summary of empirical findings on tasks of schooling and educational institutions

Educational institutions	Tasks of schooling			
	Promote equality of opportunity	Sort efficiently	Prepare for employment	Prepare for active citizenship
Differentiation	−	−	±	?
Standardization	+	+	+	?

+: Evidence points to the task benefiting from strengthening this institutional variable.
− : Evidence points to the task being impeded by strengthening this institutional variable.
±: Mixed or weak evidence regarding the relationship between institution and task.
?: Underinvestigated relationships between institution and task.

opportunity. These findings should, however, be taken as tentative findings because the study is limited to just one institutional setting.

CONCLUSIONS

We can draw several conclusions regarding the impact of national educational institutions on inequality in student achievement. In **Table 2** we return to the grid that cross-classifies educational institutions (differentiation and standardization) by four core tasks of schooling. We have now put pluses and minuses in the boxes, where an institutional feature promotes (+) or impedes (−) the realization of one of the core tasks.

A first conclusion is that equality in achievement is negatively affected by differentiation institutions. Both the variability among students and the dependence of achievement on social class and race/ethnicity are higher in educational systems that track students in different school types and school locations relative to systems that offer comprehensive schools. This effect is sizeable and is confirmed by national studies on the mechanisms underlying tracking effects. Moreover, with regard to inequality of opportunity, this finding is disputed rarely, although no consensus is found with regard to inequality as dispersion. A positive relationship is consistently found between standardization and equality of opportunity. Central examinations, especially, tend to have a positive effect on equality.

Second, the evidence with regard to efficiency in learning, operationalized by the average performance in a country, shows that educational differentiation leads to lower, rather than higher, average achievement in a number of subjects. These two conclusions imply that there is no evidence that average performance could be improved by allowing for higher inequality by means of educational differentiation. Hence, similar to the conclusions in the single-country tracking studies, the comparative literature refutes a trade-off between equality and efficiency. Thus, there is no evidence that average performance could be enhanced by allowing for a higher dispersion by means of tracking. This lack of evidence for a trade-off may result from a ceiling effect in test scores. Just like educational attainment, achievement has an upper limit, and higher averages may go together with lower dispersions as a result. With regard to standardization, it is true that average performance is increased by more standardization.

Although beyond the scope of our review, we may tentatively draw some conclusions with regard to the labor market and civic outcomes of schooling. This provides for a more comprehensive picture with regard to the institutional effects on core tasks of schooling. With regard to employment outcomes, the evidence is mixed. The effect of education on labor market outcomes tends to be stronger in strongly stratified systems owing to restricted access to tertiary education and clearer signaling (Breen 2005, Shavit & Müller 1998). Yet, to the extent that tracking decreases average performance and does not offer equal opportunity to all students, countries may not optimize their

human capital potential to the extent that would have been possible by offering comprehensive schooling. Standardization seems to reduce the noise in the signal provided by schooling for work, so there is reason to conclude that standardization enhances the labor market task of schooling.

With regard to civic skills, the evidence, unfortunately, is rather meager. Although differentiated systems have been claimed to magnify inequalities in civic skills (Terwel 2005), and single-country studies have shown differences across tracks in the kinds of civic competencies acquired (Damico et al. 1996; Niemi & Junn 1998; Ten Dam & Volman 2003, 2007), there are no studies that have compared the influence of educational institutional characteristics such as tracking on citizenship outcomes in a cross-national perspective.

The total impact of differentiation is likely larger than the comparative literature is able to show, for two reasons. First, in systems without external differentiation, forms of tracking also exist that are related to inequalities. More generally, educational systems are characterized by various sorts of segregation of students of different ability levels, social classes, and ethnic/racial groups that are not captured by national-level indicators of differentiation. The United States is a case in point: Because of the common approach of taking national-level indicators, its school system is often classified in cross-country research as comprehensive, although a vast body of American research clearly shows otherwise. It is important that future comparative research take up this issue of tracking within school types.

Second, given that externally differentiating institutions reduce the number of students eligible for college, initial inequalities in learning at mid-teen age are likely to be magnified in further educational attainment. Some studies have shown that school type has a tremendous (and realistic) effect on college expectations and aspirations in differentiating educational systems (Buchmann & Dalton 2002, Buchmann & Park 2009). Hence, it seems to be the case that differentiated educational systems channel educational careers at a very early stage, in comparison to systems without external differentiation. To gain more insight into these processes, future research may want to examine school continuation patterns conditional upon achievement in a comparative framework. Such research would shed light on how institutional factors influence the relative size of primary and secondary effects of social and ethnic/racial background. This distinction is of the utmost importance for educational policy, as reducing primary effects requires policy measures different from those of reducing secondary effects. Obviously, longitudinal school career data are needed for such an analysis and require that the authorities of cross-national student achievement data build a longitudinal design into their surveys.

Another line of research that deserves further attention is the interaction between standardization and differentiation to test the claim made by Ayalon & Gamoran (2000) that standardization can reduce the negative effects of differentiation on equality. It is plausible that standardization of educational systems diminishes the negative effects of external differentiation on inequality, as allocation to school types is based on objective and transparent criteria, reducing the effects of social origin on educational decision making.

SUMMARY POINTS

1. Countries with a more strongly differentiated (school-type tracked) educational system tend to have higher levels of inequality of educational opportunity by social class and race/ethnicity.
2. Countries with a more standardized educational system have lower levels of inequality of opportunity compared to those with unstandardized systems.

3. There is no clear pattern in the relationship between external differentiation and inequality in terms of dispersion in learning.
4. There is no evidence that average achievement is higher in countries with larger dispersions in achievement tests. This implies that there is no evidence for the existence of a trade-off between equality and efficiency in this regard.

FUTURE ISSUES

1. More research is needed on the interplay between standardization and differentiation. Does differentiation lead to higher levels of inequality in countries with low standardization?
2. Future research could also study educational institutional effects on the level and variation in social and political participation of youth and young adults. Is inequality in these outcomes larger in more strongly differentiating educational systems?

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We thank Jaap Dronkers and Sjoerd Karsten for their comments on an earlier version of this paper.

LITERATURE CITED

Alexander KL, Cook MA. 1982. Curricula and coursework: a surprise ending to a familiar story. *Am. Sociol. Rev.* 47:626–40

Alexander KL, Cook MA, McDill EL. 1978. Curriculum tracking and educational stratification: some further evidence. *Am. Sociol. Rev.* 43:47–66

Ammermüller A. 2005. *Educational opportunities and the role of institutions*. ZEW Discuss. Pap. 05-44, Cent. Eur. Econ. Res., Mannheim

Ammermüller A, Heijke H, Wössmann L. 2005. Schooling quality in Eastern Europe: educational production during transition. *Econ. Educ. Rev.* 24:579–99

Arum R, Shavit Y. 1995. Secondary vocational education and the transition from school to work. *Sociol. Educ.* 68:187–204

Atkinson AB. 1975. *The Economics of Inequality*. Oxford: Clarendon

Ayalon H, Gamoran A. 2000. Stratification in academic secondary programs and educational inequality in Israel and the United States. *Comp. Educ. Rev.* 44:54–80

Baker DP, Goesling B, LeTendre GK. 2002. Socioeconomic status, school quality, and national economic development: a cross-national analysis of the "Heyneman & Loxley effect" on mathematics and science achievement. *Comp. Educ. Rev.* 46:291–312

Bauer P, Riphahn RT. 2006. Timing of school tracking as a determinant of intergenerational transmission of education. *Econ. Lett.* 91:90–97

Becker GS. 2007. Health as human capital: synthesis and extensions. *Oxford Econ. Pap.* 59:379–410

Betts JR, Shkolnik JL. 2000. The effects of ability grouping on student achievement and resource allocation in secondary school. *Econ. Educ. Rev.* 19:1–15
Boudon R. 1974. *Education, Opportunity, and Social Inequality*. New York: Wiley
Bowles S. 2004. *Microeconomics. Behavior, Institutions, and Evolution*. Princeton, NJ: Princeton Univ. Press
Bowles S, Gintis H. 2000. Does schooling raise earnings by making people smarter? In *Meritocracy and Economic Inequality*, ed. K Arrow, S Bowles, S Durlauf, pp. 118–36. Princeton, NJ: Princeton Univ. Press
Bowles S, Gintis H. 2002. Schooling in capitalist America revisited. *Sociol. Educ.* 75:1–18
Breen R. 2005. Explaining cross-national variation in youth unemployment: market and institutional factors. *Eur. Sociol. Rev.* 21:125–34
Breen R, Jonsson JO. 2007. Explaining change in social fluidity: educational equalization and educational expansion in twentieth-century Sweden. *Am. J. Sociol.* 112:1775–810
Broaded CM. 1997. The limits and possibilities of tracking: some evidence from Taiwan. *Sociol. Educ.* 70:36–53
Brown G, Micklewright J. 2004. *Using International Surveys of Achievement and Literacy: A View from the Outside*. Montreal: UNESCO Inst. Stat.
Brown G, Micklewright J, Schnepf SV, Waldmann R. 2006. International surveys of educational achievement: how robust are the findings? *J. R. Stat. Soc. A* 170:623–46
Brunello G, Checchi D. 2007. Does school tracking affect equality of opportunity? New international evidence. *Econ. Policy* 22:781–861
Buchmann C. 2002. Measuring family background in international studies of education: conceptual issues and methodological challenges. See Porter & Gamoran 2002a, pp. 150–97
Buchmann C, Dalton B. 2002. Interpersonal influences and educational aspirations in 12 countries: the importance of institutional context. *Sociol. Educ.* 75:99–122
Buchmann C, Park H. 2009. Stratification and the formation of expectations in highly-differentiated educational systems. *Res. Soc. Stratif. Mobil.* 27:245–67
Comber LC, Keeves J. 1973. *Science Education in Nineteen Countries*. Stockholm: Almqvist & Wiksell
Crul M, Holdaway J. 2009. Children of immigrants in schools in New York and Amsterdam: the factors shaping attainment. *Teach. Coll. Rec.* 111:1476–507
Crul M, Schneider J. 2007. Integration of Turkish second-generation men and women in Germany and the Netherlands. The impact of differences in academic tracking systems. In *Educating Immigrant Youth: Pathways to Employment and Citizenship in International Perspective*, ed. M Crul, J Holdaway, C Roberts. *Spec. Issue Teach. Coll. Rec.*
Crul M, Schneider J. 2009. Children of Turkish immigrants in Germany and the Netherlands: the impact of differences in vocational and academic tracking systems. *Teach. Coll. Rec.* 111:1508–27
Crul M, Vermeulen H. 2003. The second generation in Europe. Introduction to the special issue. *Int. Migr. Rev.* 37:965–86
Damico AJ, Conway MM, Damico SB. 1996. *Democratic education: the associational life of African-American and white high school students*. Presented at Annu. Meet. Am. Educ. Res. Assoc., April, New York City
Darling-Hammond L. 1994. Performance-based assessment and educational equity. *Harvard Educ. Rev.* 64:5–30
Dee TS. 2004. Are there civic returns to education? *J. Public Econ.* 88:1697–720
Dobbelsteen S, Levin J, Oosterbeek H. 2002. The causal effect of class size on scholastic achievement: distinguishing the pure class size effect from the effect of changes in class composition. *Oxford Bull. Econ. Stat.* 64:17–38
Dronkers J, Robert P. 2008. Differences in scholastic achievement of public, private government-dependent, and private independent schools. A cross-national analysis. *Educ. Policy* 22:541–77
Duru-Bellat M, Kieffer A. 2000. Inequalities in educational opportunities in France: educational expansion, democratization or shifting barriers? *J. Educ. Policy* 15:333–52
Duru-Bellat M, Suchaut B. 2005. Organisation and context, efficiency and equity of educational systems: what PISA tells us. *Eur. Educ. Res. J.* 4:181–94
Entorf H, Lauk M. 2008. Peer effects, social multipliers and migrants at school: an international comparison. *J. Ethnic Migr. Stud.* 34:633–54
Erikson R, Goldthorpe JH, Jackson M, Yaish M, Cox DR. 2005. On class differentials in educational attainment. *Proc. Natl. Acad. Sci. USA* 102:9730–33

Figlio DN, Page ME. 2002. School choice and the distributional effects of ability tracking: Does separation increase inequality? *J. Urban Econ.* 51:497–514
Fuchs T, Wössmann L. 2007. What accounts for international differences in student performance? A re-examination using PISA data. *Empir. Econ.* 32:433–64
Gamoran A. 1992. The variable effects of high school tracking. *Am. Sociol. Rev.* 57:812–28
Gamoran A. 1996. Curriculum standardization and equality of opportunity in Scottish secondary education: 1984–90. *Sociol. Educ.* 69:1–21
Gamoran A, Berends M. 1987. The effects of stratification in secondary schools: synthesis of survey and ethnographic research. *Rev. Educ. Res.* 57:415–35
Gamoran A, Mare RD. 1989. Secondary school tracking and educational inequality: compensation, reinforcement, or neutrality? *Am. J. Sociol.* 94:1146–83
Gamoran A, Weinstein M. 1998. Differentiation and opportunity in restructured schools. *Am. J. Educ.* 106:385–415
Goldstein H. 2004. International comparisons of student attainment: some issues arising from the PISA study. *Assess. Educ.* 11:319–30
Goldthorpe JH. 1996. Class analysis and the reorientation of class theory: the case of persisting differentials in educational attainment. *Br. J. Sociol.* 47:481–505
Hallinan MT. 1988. Equality of educational opportunity. *Annu. Rev. Sociol.* 14:249–68
Hallinan MT. 1991. School differences in tracking structures and track assignments. *J. Res. Adolesc.* 1:251–75
Hallinan MT. 1992. Organization of students for instruction in the middle school. *Sociol. Educ.* 65:114–27
Hallinan MT. 1994a. School differences in tracking effects on achievement. *Soc. Forces* 72:799–820
Hallinan MT. 1994b. Tracking: from theory to practice. *Sociol. Educ.* 67:79–84
Hallinan MT. 1996a. Race effects on students' track mobility in high school. *Soc. Psychol. Educ.* 1:1–24
Hallinan MT. 1996b. Track mobility in secondary school. *Soc. Forces* 74:983–1002
Hallinan MT, Kubitschek WN. 1999. Curriculum differentiation and high school achievement. *Soc. Psychol. Educ.* 3:41–62
Hanushek EA. 2006. Alternative school policies and the benefits of general cognitive skills. *Econ. Educ. Rev.* 25:447–62
Hanushek EA, Jamison DT, Jamison EA, Wossmann L. 2008. Education and economic growth: It's not just going to school, but learning something while there that matters. *Educ. Next* 8:62–70
Hanushek EA, Wösmann L. 2005. Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries. *Econ. J.* 116:C63–76
Hattie JAC. 2002. Classroom composition and peer effects. *Int. J. Educ. Res.* 35:449–81
Hauser RM, Featherman DL. 1976. Equality of schooling: trends and prospects. *Sociol. Educ.* 49:99–120
Heckman JJ. 2000. Policies to foster human capital. *Res. Econ.* 54:3–56
Heyneman SP, Loxley WA. 1983. The effect of primary-school quality on academic achievement across twenty-nine high and low-income countries. *Am. J. Sociol.* 88:1162–94
Hillygus DS. 2005. The missing link: exploring the relationship between higher education and political engagement. *Polit. Behav.* 27:25–47
Hopper EI. 1968. A typology for the classification of educational systems. *Sociology* 2:29–46
Horn D. 2009. Age of selection counts: a cross-country analysis of educational institutions. *Educ. Res. Eval.* 15:343–66
Huang M-H. 2009. Classroom homogeneity and the distribution of student math performance: a country-level fixed-effects analysis. *Soc. Sci. Res.* 38:781–91
Husen T, ed. 1967. *International Study of Achievement in Mathematics. A Comparison of Twelve Countries*. Stockholm: Almqvist & Wiksell
Husen T. 1973. The standard of the elite: some findings from the IEA international survey in mathematics and science. *Acta Sociol.* 16:305–23
Kalmijn M, Kraaykamp G. 2003. Dropout and downward mobility in the educational career: an event-history analysis of ethnic schooling differences in the Netherlands. *Educ. Res. Eval.* 9:265–87
Kenworthy L. 2008. *Jobs with Equality*. New York: Oxford Univ. Press
Kerckhoff AC. 2001. Education and social stratification processes in comparative perspective. *Sociol. Educ.* Extra Issue:3–18

Kirsch I, De Jong J, Lafontaine D, McQueen J, Mendelovits J, Monseur C. 2002. *Reading for Change: Performance and Engagement Across Countries: Results from PISA 2000*. Paris: OECD
Kubitschek WN, Hallinan MT. 1998. Tracking and students' friendships. *Soc. Psychol. Q.* 61:1–15
Kulik C-L, Kulik JA. 1982. Effects of ability grouping on secondary school students: a meta-analysis of evaluation findings. *Am. Educ. Res. J.* 19:415–28
LeTendre GK, Hofer BK, Shimizu H. 2003. What is tracking? Cultural expectations in the United States, Germany, and Japan. *Am. Educ. Res. J.* 40:43–89
Lucas SR. 1999. *Tracking Inequality. Stratification and Mobility in American High Schools*. New York: Teach. Coll. Press
Lucas SR. 2001. Effectively maintained inequality: education transitions, track mobility, and social background effects. *Am. J. Sociol.* 106:1642–90
Marginson S, Weko T, Channon N, Luukkonen T, Oberg J. 2007. *Thematic Review of Tertiary Education*. Paris: OECD
Marks GN. 2005. Cross-national differences and accounting for social class inequalities in education. *Int. Sociol.* 20:483–505
Marsh HW. 1984. Self-concept, social comparison, and ability grouping: a reply to Kulik and Kulik. *Am. Educ. Res. J.* 21:799–806
Mayer T. 1960. The distribution of ability and earnings. *Rev. Econ. Stat.* 42:189–95
Meghir C, Palme M. 2005. Educational reform, ability, and family background. *Am. Econ. Rev.* 95:414–24
Michaelowa K, Bourdon J. 2006. *The impact of student diversity in secondary schools: an analysis of the international PISA data and implications for the German education system*. HWWI Res. Pap. 3–2, Hamburg Inst. Int. Econ., Hamburg
Micklewright J, Schnepf SV. 2007. Inequality of learning in industrialized countries. In *Inequality and Poverty Re-Examined*, ed. SP Jenkins, J Micklewright, pp. 129–45. Oxford: Oxford Univ. Press
Miller D. 1999. *Principles of Social Justice*. Cambridge, MA: Harvard Univ. Press
Muller C, Schiller KS. 2000. Leveling the playing field? Students' educational attainment and states' performance testing. *Sociol. Educ.* 73:196–218
Müller W, Gangl M, eds. 2003. *Transitions From Education to Work in Europe—the Integration of Youth into EU Labor Markets*. Oxford: Oxford Univ. Press
Niemi RG, Junn J. 1998. *Civic Education. What Makes Students Learn?* New Haven, CT: Yale Univ. Press
Oakes J. 1990. Opportunities, achievement and choice: women and minority students in science and mathematics. *Rev. Res. Educ.* 16:153–222
Oakes J. 2005 [1985]. *Keeping Track. How Schools Structure Inequality*. New Haven, CT: Yale Univ. Press
OECD. 2004. *Learning for Tomorrow's World: First Results from PISA 2003*. Paris: OECD
OECD. 2005. PISA 2003. *Tech. Rep., Org. Econ. Coop. Dev.*, Paris: OECD
OECD. 2008. *Tertiary Education for the Knowledge Society*. Vol. II. Paris: OECD
Pallas AM, Entwistle DR, Alexander KL, Stukla FM. 1994. Ability-group effects: instructional, social, or institutional? *Sociol. Educ.* 67:27–46
Park H. 2008a. Home literacy environments and children's reading performance: a comparative study of 25 countries. *Educ. Res. Eval.* 14:489–505
Park H. 2008b. The varied educational effects of parent-child communication: a comparative study of fourteen countries. *Comp. Educ. Rev.* 52:219–43
Paulle B. 2005. *Anxiety and Intimidation in the Bronx and the Bijlmer: An Ethnographic Comparison of Two Schools*. Amsterdam: Dutch Univ. Press
Pekkarinen T, Uusitalo R, Kerr S. 2009. *School tracking and development of cognitive skills*. IZA Discuss. Pap. 4058, Inst. Study Labor, Bonn
Pfeffer FT. 2008. Persistent inequality in educational attainment and its institutional context. *Eur. Sociol. Rev.* 24:543–65
Porter AC, Gamoran A. 2002a. *Methodological Advances in Cross-National Surveys of Educational Achievement*. Washington, DC: Natl. Acad. Press
Porter AC, Gamoran A. 2002b. Progress and challenges for large-scale studies. See Porter & Gamoran 2002a, pp. 3–23

Ramirez FO, Luo X, Schofer E, Meyer JW. 2006. Student achievement and national economic growth. *Am. J. Educ.* 113:1–29
Raudenbush SW, Kim J-S. 2002. Statistical issues in analysis of international comparisons of educational achievement. See Porter & Gamoran 2002a, pp. 267–94
Rijken S. 1999. *Educational expansion and status attainment. A cross-national and over-time comparison*. ICS Diss. Univ. Utrecht
Rindermann H. 2007. The *g*-factor of international cognitive ability comparisons: the homogeneity of results in PISA, TIMSS, PIRLS and IQ tests across nations. *Eur. J. Personal.* 21:667–706
Schütz G, Ursprung H, Wössmann L. 2008. Education policy and equality of opportunity. *Kyklos* 61:279–308
Shavit Y. 1984. Tracking and ethnicity in Israeli secondary education. *Am. Sociol. Rev.* 49:210–20
Shavit Y. 1990. Segregation, tracking, and the educational attainment of minorities: Arabs and oriental Jews in Israel. *Am. Sociol. Rev.* 55:115–26
Shavit Y, Müller W, eds. 1998. *From School to Work. A Comparative Study of Educational Qualifications and Occupational Destinations*. Oxford: Clarendon
Stanat P, Christensen G. 2006. *Where Immigrant Students Succeed: A Comparative Review of Performance and Engagement in PISA 2003*. Paris: OECD
Stevenson DL, Baker David P. 1991. State control of the curriculum and classroom instruction. *Sociol. Educ.* 64:1–10
Ten Dam GTM, Volman MLL. 2003. A life jacket or an art of living: inequality in social competence education. *Curric. Inq.* 33:117–37
Ten Dam GTM, Volman MLL. 2007. Educating for adulthood or for citizenship: social competence as an educational goal. *Eur. J. Educ.* 42:281–98
Terwel J. 2005. Curriculum differentiation: multiple perspectives and developments in education. *J. Curric. Stud.* 37:653–70
Thomas V, Wang Y, Fan X. 2001. *Measuring Education Inequality: Gini Coefficients of Education*. Washington, DC: World Bank
Thrupp M, Lauder H, Robinson T. 2002. School composition and peer effects. *Int. J. Educ. Res.* 37:483–504
Torney-Purta J, Lehmann R, Oswald H, Schulz W. 2001. *Citizenship and Education in Twenty-Eight Countries. Civic Knowledge and Engagement at Age Fourteen*. Amsterdam: Int. Assoc. Eval. Educ. Achiev.
Turner RH. 1960. Sponsored and contest mobility and the school system. *Am. Sociol. Rev.* 25:855–67
Van de Werfhorst HG, Van Tubergen F. 2007. Ethnicity, schooling, and merit in the Netherlands. *Ethnicities* 7:416–44
Van den Eeden P, Terwel J. 1994. Evaluation of a mathematics curriculum: differential effects. *Stud. Educ. Eval.* 20:457–75
Vandenberge V. 2006. Achievement effectiveness and equity: the role of tracking, grade repetition and inter-school segregation. *Appl. Econ. Lett.* 13:685–93
Van Elk R, van der Steeg M, Webbink D. 2009. *The effect of early tracking on participation in higher education*. CPB Neth. Bur. Econ. Policy Anal. No. 182, The Hague
Walzer M. 1983. *Spheres of Justice*. New York: Basic Books
Westbury I, Hsu C-S. 1996. Structures of curriculum governance and classroom practice in mathematics. *Educ. Eval. Policy Anal.* 18:123–39
Willms JD. 1986. Social class segregation and its relationship to pupils' examination results in Scotland. *Am. Sociol. Rev.* 51:224–41
Wolf A. 2002. *Does Education Matter? Myths about Education and Economic Growth*. London: Penguin
Wolff EN. 2006. *Does Education Really Help? Skill, Work and Inequality*. New York: Oxford Univ. Press
Wössmann L. 2003a. Central exit exams and student achievement: international evidence. In *No Child Left Behind? The Politics and Practice of Accountability*, ed. PE Peterson, MR West, pp. 292–324. Washington, DC: Brookings Inst. Press
Wössmann L. 2003b. Schooling resources, educational institutions and student performance: the international evidence. *Oxford Bull. Econ. Stat.* 65:117–70
Wössmann L. 2005. The effect heterogeneity of central examinations: evidence from TIMSS, TIMSS-repeat and PISA. *Educ. Econ.* 13:143–69

Wössmann L. 2008b. Efficiency and equity of European education and training policies. *Int. Tax Public Financ.* 15:199–230

Wössmann L. 2008a. How equal are educational opportunities? Family background and student achievement in Europe and the US. *Z. Betriebswirtsch.* 78:45–70

Yamamoto K, Kulick E. 2000. Scaling methodology and procedures for the TIMSS mathematics and science scales. In *TIMSS 1999 Tech. Rep.*, ed. MO Martin, KD Gregory, SE Stemler, pp. 237–63. Chestnut Hill, MA: Int. Study Cent. Boston Coll.

Zimmer RW, Toma EF. 2000. Peer effects in private and public schools across countries. *J. Policy Anal. Manag.* 19:75–92

Annual Review
of Sociology
Volume 36, 2010

Contents

Institutions and Culture

Formal Organizations

Political and Economic Sociology

Differentiation and Stratification

Individual and Society

Policy

Indexes

Errata

An online log of corrections to *Annual Review of Sociology* articles may be found at http://soc.annualreviews.org/errata.shtml