

## The impact of education quality on development goals

It is commonly presumed that formal schooling is one of several important contributors to the skills of an individual and to human capital. It is not the only factor. Parents, individual abilities and friends undoubtedly contribute. Schools nonetheless have a special place, not only because education and 'skill creation' are among their prime explicit objectives, but also because they are the factor most directly affected by public policies. It is well established that the distribution of personal incomes in society is strongly related to the amount of education people have had. Generally speaking more schooling means higher lifetime incomes. These outcomes emerge over the long term. It is not people's income while in school that is affected, nor their income in their first job, but their income over the course of their working life. Thus, any noticeable effects of the current quality of schooling on the distribution of skills and income will become apparent some years in the future, when those now in school become a significant part of the labour force.

### Impact of quality on individual incomes

One challenge in documenting the impact of differences in the quality of human capital has been its measurement. Much of the discussion of quality – in part related to new efforts to improve accountability – has identified the importance of

enhancing cognitive skills via schooling, and most parents and policy makers accept that such skills represent a key dimension of schooling outcomes. If cognitive skills do provide proxy evidence, however incomplete, for school quality, the question arises as to whether these skills are correlated with students' subsequent performance in the labour market and with the economy's ability to grow.

There is mounting evidence that the quality of human resources, as measured by test scores, is directly related to individual earnings, productivity and economic growth. A range of research results from the United States shows that the earnings advantages due to higher achievement on standardized tests are quite substantial.<sup>2</sup> These studies typically find that measured achievement has a clear impact on earnings, after allowing for differences in the quantity of schooling, age or work experience, and for other factors that might influence earnings. In other words, for those leaving school at a given grade, higher-quality school outcomes (represented by test scores) are closely related to subsequent earnings differences and, we therefore suppose, to differences in individual productivity.

Three recent studies from the United States provide direct and quite consistent estimates of the impact of test performance on earnings (Mulligan, 1999; Murnane et al., 2000; Lazear, 2003). They use different data sets – each of them nationally representative – following students after they leave school and enter the labour force. They suggest that one standard deviation increase in mathematics performance at the end of high school translates into 12% higher annual earnings.<sup>3</sup> By way of comparison, estimates of the average value of an additional year of school attainment in the United States are typically 7–10%.

There are reasons to believe that these estimates provide a lower boundary for the impact of higher cognitive achievement on earnings. First, they are obtained fairly early in the working lives of the sampled people, who were generally 25 to 35 years old at the dates to which the data refer, and evidence suggests that the impact of test performance increases with work experience.<sup>4</sup> Second, the observed labour market experiences cover 1985–95, and other evidence suggests that the value of skills and schooling has grown since

**It is well established that the distribution of personal incomes in society is strongly related to the amount of education people have had**

1. The existing literature, whether in economics or in education science, has focused on educational outcomes rather than inputs and processes, and indeed on one type of outcome only: cognitive skills. Accordingly, most of this chapter focuses on cognitive achievement, though it also mentions the importance of non-cognitive skills and other outcomes of schooling whose value is increasingly recognised.

2. These results are derived from different approaches, but the underlying analysis involves estimating a standard Mincer earnings function and adding a measure of individual cognitive skills. This approach relates the logarithm of earnings to years of schooling, experience and other factors that might yield individual earnings differences. The clearest analyses are found in Bishop (1989, 1991), O'Neill (1990), Grogger and Eide (1993), Blackburn and Neumark (1993, 1995), Murnane, Willett and Levy (1995), Neal and Johnson (1996), Mulligan (1999), Murnane et al. (2000), Altonji and Pierret (2001), Murnane et al. (2001) and Lazear (2003).

3. One standard deviation increase from the mean would be an achievement level equivalent to the eighty-fifth percentile of the distribution; i.e. 15% of students would normally achieve higher test scores than this. Murnane et al. (2000) provide evidence from the High School and Beyond study and the National Longitudinal Survey of the High School Class of 1972. Their estimates suggest some variation, with males obtaining a 15% increase and females a 10% increase per standard deviation of test performance. Lazear (2003), relying on a somewhat younger sample from the National Education Longitudinal Study of 1988, provides a single estimate of 12%. Similarly, Mulligan (1999) finds 11% for the normalized Armed Forces Qualification Test score in the National Longitudinal Survey of Youth data.

4. Altonji and Pierret (2001) find that the impact of achievement on earnings grows with experience partly because the employer has more chance to observe performance.

then. Third, future general improvements in productivity throughout the economy are likely to lead to larger returns to higher skill levels.<sup>5</sup>

As regards other direct benefits, research has established strong returns to both numeracy and literacy in the United Kingdom<sup>6</sup> and to literacy in Canada.<sup>7</sup> Accordingly, educational programmes that deliver these skills will bring higher individual economic benefits than those that do not.

Part of the returns to school quality comes through continuation in school.<sup>8</sup> Obviously, students who do better in school, as evidenced by either examination grades or scores on standardized achievement tests, tend to go further in school or university.<sup>9</sup> By the same token, the net costs of improvements in school quality, if reflected in increased attainment by learners, are less than they appear – perhaps substantially so – because of the resulting reductions in rates of repetition and dropout. Thus, higher student achievement keeps students in school longer, which leads, among other things, to higher completion rates at all levels of schooling. Accordingly, in countries where schools are dysfunctional and grade repetition is high, some improvements in quality may be largely self-financing, by reducing the average time completers spend in school.

As regards these relationships in developing countries, it appears likely, on the basis of somewhat limited evidence, that the returns to school quality are, if anything, higher than in more industrialized contexts. Table 2.1 provides a simple summary of research results for six countries, mainly in Africa. Using simple measures of basic cognitive skills, these studies show that such skills are separately important in determining earnings, apart from the effect of years of schooling attained. Although there are reasons for caution in interpreting the results,<sup>10</sup> the table suggests the presence of strong economic returns to education quality. Only the studies for Ghana and the United Republic of Tanzania had ranges of returns that were less than or similar to the United States estimates. Elsewhere, one standard deviation increase in test scores was associated with wage increases ranging from 12% to 48%, suggesting a substantial return to higher levels of cognitive skills and probably, therefore, to higher levels of school quality.

## Impact of quality on economic growth

The relationship between measured labour force quality and economic growth is perhaps even more important than the impact of human capital and school quality on individual productivity and incomes. Economic growth determines how much improvement can occur in the overall standard of living of a society. Moreover, the education of each individual has the possibility of making others better off (in addition to the individual benefits just discussed). Specifically, a more educated society may translate into higher rates of innovation, higher overall productivity through firms' ability to introduce new and better production methods, and faster introduction of new technology. These externalities provide extra reason for being concerned about the quality of schooling.

Economists have developed a variety of models and ideas to explain differences in growth rates among countries, invariably featuring the

**A more educated society may translate into higher rates of innovation, higher overall productivity and faster introduction of new technology**

5. Studies on the impact of achievement on earnings typically compare workers of different ages at one point in time, in order to obtain an estimate of how earnings will change for any individual. Any productivity improvements in the economy, however, will tend to raise the earnings of individuals over time. Thus, the benefits of improvements in student skills are likely to grow over a person's working life, rather than remain constant.

6. See McIntosh and Vignoles (2001). Because they look at discrete levels of skills, it is difficult to compare the quantitative magnitudes directly with the United States work.

7. Finnie and Meng (2002) and Green and Riddell (2003) both suggest that literacy has a significant return, but Finnie and Meng find an insignificant return to numeracy, a finding at odds with most other analyses focusing on numeracy or mathematics skills.

8. Much of the work by economists on differences in worker skills has been directed at determining the average labour market returns to additional schooling. The argument has been that, as higher-ability students are more likely to continue in schooling, part of the higher earnings observed for those with additional schooling really reflects pay for added ability rather than additional schooling. Economists have pursued a variety of analytical approaches for dealing with this, including adjusting for measured cognitive test scores, but this work generally ignores issues of variation in school quality. The approaches have included looking for circumstances where the amount of schooling is affected by things other than the student's valuation of continuing, and considering the income differences among twins (see Card, 1999). The various adjustments for ability differences typically result in small changes to the estimates of the value of schooling, and Heckman and Vytlačil (2001) argue that it is not possible to separate the effects of ability and schooling. The only explicit consideration of school quality typically investigates expenditure and resource differences among schools, but these are known to be poor measures of school quality differences (Hanushek, 2002a).

9. Though the point may indeed be obvious, a significant amount of research evidence also documents it. See, for example, Dugan (1976) and Manski and Wise (1983). Rivkin (1995) finds that variations in test scores in the USA capture a considerable proportion of the systematic variation in high school completion and college continuation. Bishop (1991) and Hanushek, Rivkin and Taylor (1996), in considering the factors that influence school attainment, find that individual achievement scores are highly correlated with continued school attendance. Behrman et al. (1998) find strong achievement effects on both continuation into college and college quality; moreover, the effects are greater when proper account is taken of the various determinants of achievement. Hanushek and Pace (1995) find that college completion is significantly related to higher test scores at the end of high school.

10. The estimates appear to be quite sensitive to the estimation methodology. Both within individual studies and across studies using the same basic data, the results are quite sensitive to the techniques employed in revealing the fundamental parameter for cognitive skills. See Glewwe (2002).

**Table 2.1: Estimated returns to a standard deviation increase in cognitive skills**

Study	Country	Estimated effect <sup>1</sup>	Notes
Glewwe (1996)	Ghana	0.21** to 0.3** (government) 0.14 to 0.17 (private)	Alternative estimation approaches yield some differences; mathematics effects shown to be generally more important than reading effects, and all hold even with Raven's test for ability.
Jolliffe (1998)	Ghana	0.05 to 0.07*	Household income related to average mathematics score with relatively small variation by estimation approach; effect from off-farm income with on-farm income unrelated to skills.
Vijverberg (1999)	Ghana	uncertain	Income estimates for mathematics and reading with non-farm self-employment; highly variable estimates (including both positive and negative effects) but effects not generally statistically significant.
Boissiere, Knight and Sabot (1985); Knight and Sabot (1990)	Kenya	0.19** to 0.22**	Total sample estimates: small variation by primary and secondary school leavers.
Angrist and Lavy (1997)	Morocco	uncertain	Cannot convert to standardized scores because use indexes of performance; French writing skills appear most important for earnings, but results depend on estimation approach.
Alderman et al. (1996)	Pakistan	0.12 to 0.28*	Variation by alternative approaches and by controls for ability and health; larger and more significant without ability and health controls.
Behrman, Ross and Sabot (forthcoming)	Pakistan	uncertain	Estimates of structural model with combined scores for cognitive skill; index significant at .01 level but cannot translate directly into estimated effect size.
Moll (1998)	South Africa	0.34** to 0.48**	Depending on estimation method, varying impact of computation; comprehension (not shown) generally insignificant.
Boissiere, Knight and Sabot (1985); Knight and Sabot (1990)	UR Tanzania	0.07 to 0.13*	Total sample estimates: smaller for primary than secondary school leavers.

Notes: \*significant at .05 level; \*\*significant at .01 level.

1. Estimates indicate proportional increase in wages from an increase of one standard deviation in measured test scores.

Source: Hanushek (2004)

11. For a review of analyses and of the range of factors they include, see Barro and Sala-i-Martin (2003).

12. See also Barro and Lee (2001), whose analysis of qualitative differences includes literacy.

13. For details of this work see Hanushek and Kimko (2000) and Hanushek (2003b). Significantly, adding other factors potentially related to growth, including aspects of international trade, private and public investment and political instability, leaves the effects of labour force quality unchanged. The results also prove robust after allowing for other factors that can cause both higher growth and better educational performance.

14. Other desirable outcomes, apart from those relating to the competence of the labour force, that stem from improvements in education quality are discussed below.

importance of human capital.<sup>11</sup> In testing these models, empirical work has emphasized school attainment differences as a proxy for differences in human capital. Many studies find that the quantity of schooling, measured this way, is closely related to economic growth rates. The quantity of schooling, however, is a very crude measure of knowledge and cognitive skills – particularly in an international context, where wide differences exist as regards the resources available to school systems and the levels of household poverty.

Difficulties in international comparison of education quality have hampered attempts to incorporate measures of the quality of schooling in empirical analyses. In recent years, however, the existence of international achievement tests, administered in a consistent way to a growing group of countries, has begun to make such comparison possible. Hanushek and Kimko

(2000), for example, incorporate information about international differences in mathematics and science knowledge by developing a common scale across all countries and tests and including a composite measure of quality as an additional determining variable in cross-country growth equations.<sup>12</sup> Their results suggest a strong impact of differences in school quality on economic growth: a difference of one standard deviation on test performance is related to a 1% difference in annual growth rates of GDP per capita.<sup>13</sup> That may sound small, but it is actually very significant. Because the added growth has a compound effect, it brings powerful incremental results for national income and societal well-being. Thus, the quality of the labour force, as measured by mathematics and science scores, appears to be an important determinant of growth, and thus of the potential to alleviate poverty.<sup>14</sup>

## Quality and non-cognitive skills

There is a whole set of non-cognitive skills that are important for success in economic life. As Aesop's fable of the Tortoise and the Hare sets out to demonstrate, those with motivation and perseverance are likely to do better, other things being equal, than people of similar intelligence but less staying power. It has become increasingly clear that society rewards these and other non-cognitive skills such as honesty, reliability, determination and personal efficacy.

Early research found that personality and behavioural traits such as perseverance and leadership qualities had a significant influence upon labour market success, including earnings (Jencks et al., 1979). Personal stability, dependability, willingness to adopt the norms of institutions and hierarchies – these were shown to be important conditions for getting on in life and winning employer approval (Bowles and Gintis, 1976). Until recently, data and measurement problems largely discouraged further attempts to estimate the effects of such characteristics. However, a recent study of United States and United Kingdom data finds that individual differences in personality account for substantial differences in earnings, and that the way such characteristics affect earnings differs between the sexes (Bowles, Gintis and Osborne, 2001). In high-status jobs, women are penalized for having aggressive personalities, whereas men are rewarded, the study finds (after controlling for education, measured ability, exam success and other factors affecting earnings). The pattern is reversed for passive, withdrawing personalities, with men losing and women gaining income. The study also finds, again after controlling for other income-related factors, that women in the United States with a lower sense of their own ability to influence their destinies have lower earnings. Other recent research from the United States shows that bright but undisciplined male school dropouts who lack persistence and adaptability earn less than others with the same levels of ability and cognitive achievement and will continue to do so, beyond school (Heckman and Rubenstein, 2001). These types of enquiry are increasingly demonstrating the importance of non-cognitive skills in economic life.

Such skills are imparted and nourished by schools, at least in part. Not all are necessarily

desirable; some (honesty, determination, reliability) are encouraged and rewarded by schools while other non-cognitive traits that the labour market appears to value (passivity in women, aggressiveness in men) are targeted by many schools as undesirable outcomes that strengthen inequalities in society. On average, the possession of useful non-cognitive skills may be approximated by test scores, in that higher cognitive achievers may have more of these 'valuable' non-cognitive skills too. But it is likely that their distribution explains some of the variation in earnings among those with similar cognitive achievement levels, indicating that these skills and traits are separately valued in the labour market.

## The impact of quality on behavioural change

It seems, then, that there is good evidence to suggest that the quality of education – as measured by test scores – has an influence upon the speed with which societies can become richer and the extent to which individuals can improve their own productivity and incomes. We also know that years of education and acquisition of cognitive skills – particularly the core skills of literacy and numeracy – have economic and social pay-offs as regards income enhancement, improved productivity in both rural non-farm and urban environments and strengthened efficacy of household behaviour and family life (Jolliffe, 1998; Rosenzweig, 1995). In South Africa and Ghana, the number of years spent at school is negatively correlated with fertility rates, a relationship partly deriving from links between cognitive achievement and fertility (Thomas, 1999; Oliver, 1999).<sup>15</sup> Education systems that are more effective in establishing cognitive skills to an advanced level and distributing them broadly through the population will bring stronger social and economic benefits than less effective systems. This implies that the subject structure of the curriculum is important, in that school systems that do not impart literacy and numeracy would not be associated with these benefits – and those that do so more effectively (i.e. those that are of higher quality) are associated with larger benefits.

Clearly, then, differences in education quality can affect human behaviour in ways that facilitate the achievement of a wide range of human goals.

**In high-status jobs, women are penalized for having aggressive personalities, whereas men are rewarded**

15. The exceedingly complex links between education and fertility have been researched for many years. It is not only cognitive skills but also the process of socialization through schooling that can help give women the autonomy to change fertility outcomes (see Basu, 2002).

**The cognitive skills required for informed choices in respect of HIV/AIDS risk appear to be based on levels of education and literacy**

Granted, knowledge, even when widely shared, is not sufficient in and of itself to change behaviour. Opportunities of many kinds, however, can be found to improve the quality of schooling so as to facilitate such consequences. One important current example concerns health behaviour – specifically the challenge of responding to the HIV/AIDS pandemic.<sup>16</sup>

The mounting evidence of HIV/AIDS' impact in many countries indicates the potential importance of links between HIV/AIDS education and behavioural change. We readily and reasonably assume that the provision of clear information about the sources of HIV/AIDS infection and, indeed, improved general levels of literacy, will allow those at risk to understand and judge their options better. Are we right to do so? Box 2.1 indicates that knowledge and risk-reducing skills are acquired through a complex network of formal and informal sources, of which the education system is only one. Nevertheless, the cognitive skills required for informed choices in respect of HIV/AIDS risk – and for behavioural change – appear to be substantively based on levels of education and literacy. Thus, the primary inherent value of formal education in this context is to enhance the learning skills required to understand the HIV/AIDS education on offer and make sense of the many related messages from other sources (Badcock-Walters, Kelly and Gørgens, 2004). This suggests that access to and retention in the school system is indeed the uniquely important 'social vaccine' to which many refer (Kelly, 2000; Low-Beer and Stoneburner, 2001). Helping schools deliver effective messages about HIV/AIDS prevention can only enhance their beneficial impact.

## International assessments of cognitive achievement

In much of the evidence on the relationships between education quality and levels of economic growth and personal incomes, reviewed earlier, test scores serve as a proxy for education quality. Assessment of learners' progress, using cognitive tests, serves a number of purposes. It can provide an indication of how well items in the curriculum are being learned and understood, for example – a 'formative' influence for teaching and learning policies at local or national level. Equally, it can provide a signal as to how well

learners have done at the main exit points from the school system, thereby typically helping educational institutions or employers to select those best qualified for further education or for various kinds of work. This type of 'summative' assessment is used as a means of facilitating (and legitimizing) access to social and economic hierarchies. Precisely because of their role in rationing access to scarce opportunities, such assessments can have an important impact on what goes on in schools. They may have beneficial effects by helping to ensure that the intended curriculum is taught and learned, but they can bring unintended, detrimental effects where the pressure to succeed encourages excessive attention to passing examinations rather than to broader aspects of learning.

These and other aspects of national educational assessment systems, and the impact they can have upon the quality of education, are discussed further in Chapter 4. Here we are interested in the large – and growing – body of information available from international surveys of cognitive achievement, upon which most international comparisons of education quality draw. What can their results tell us about the determinants of education quality?

### The studies

In the late 1950s, the International Association for the Evaluation of Educational Achievement (IEA) was formed. It initiated what would become a major set of studies aiming to measure cognitive achievement at various levels of education in several countries and to identify the main causes of differences in outcomes. Twelve countries joined its first mathematics study. By 2000, some fifty countries were participating in surveys covering mathematics and science (now called the Trends in International Mathematics and Science Study or TIMSS), science, reading (the Progress in International Reading Literacy Study or PIRLS) and other subjects. Strongly influenced by the IEA experience, several other such studies, usually of regional focus, have since been established. They include the Programme for International Student Assessment (PISA), set up by the OECD in 1998 and now covering fifty-nine mainly industrialized and middle-income countries; the Southern and Eastern African Consortium for Monitoring

16. A second example is the impact of educational change on gender relations in school and in society. It is clear that changes in school location planning, reforms to curricula and textbook development, widening subject options for girls, changing the nature of school chores, improving teacher training and sensitization, ensuring that school facilities are girl-friendly, making timetables more flexible to respond to the demands of households, and a wide range of other, more detailed reforms can help reduce gender inequality in school and beyond. These matters comprised the major theme of the *EFA Global Monitoring Report 2003/4* (UNESCO, 2003a). See that volume for extensive discussion and evidence on these issues.

**Box 2.1 Education and HIV/AIDS risk avoidance: does knowledge equal change?**

A recent study of six African countries - Kenya, Malawi, Uganda, the U. R. Tanzania, Zambia and Zimbabwe - found high levels of HIV/AIDS awareness in the population (more than 90% of those surveyed) in each country but considerable differences as to the sources of this knowledge (Low-Beer and Stoneburner, 2000). Social networks of friends and family were the main sources in Uganda, whereas in the other countries the mass media and institutional sources (schools, churches, clinics) predominated.

The percentage of respondents who had known someone with HIV/AIDS was substantially higher in Uganda (91.5% of men and 86.4% of women) than in the five other countries. This direct experience appears to have acted as a spur to behavioural change. For example, about 20% of Ugandan men aged 15 to 24 who knew someone with AIDS had started using condoms, whereas only some 5% of those who did not know an AIDS sufferer used them.

In a South African study, almost one-fifth of 15- to 24-year-olds indicated that they talk to teachers and classmates about HIV/AIDS, and about one-third of them reported learning most about HIV/AIDS from school sources (Pettifor et al., 2004). On the other hand, among secondary school students in Botswana, Malawi and Uganda, radio was the most widely cited source (Bennell, Hyde and Swainson, 2002). Teachers ranked second in Botswana and Malawi for both genders, yet in Uganda they were ranked fifth by male students and second by female students, reinforcing the view that education in schools is not necessarily the principal source of information about HIV/AIDS in that country.

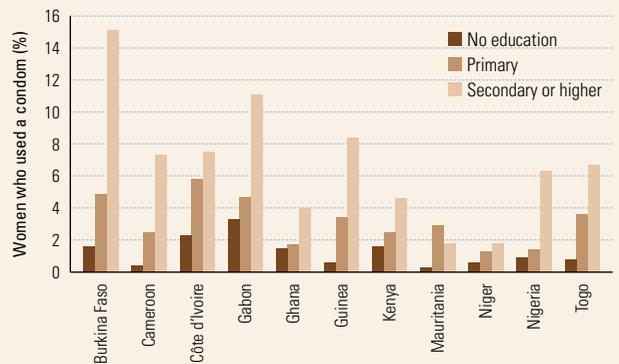
Even in Uganda, though, clear evidence exists of strong and increasing links between HIV/AIDS education, increased general knowledge and risk-avoidance behaviour. Figure 2.1 shows that rates of HIV prevalence in rural Uganda were initially closely comparable for all education levels, but separation began in 1995, and by the turn of the century those with some secondary education had much lower prevalence rates than those with less schooling. This evidence is mirrored in other African countries, where condom use is rising sharply among both men and women with higher levels of schooling (Figures 2.2 and 2.3). Thus, retention in a functional education system seems likely to provide the quality of education and skills development necessary to reduce or eliminate sexual and lifestyle risk. The general cognitive and social gains from a basic education seem to be the main factor in protecting adolescents and young adults from infection.

**Figure 2.1: HIV prevalence in rural Uganda (%) by education category, 1990-2001 (individuals aged 18-29)**



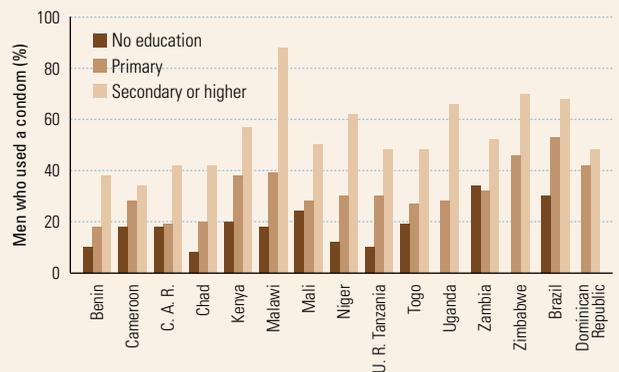
Note: Primary means having attended any or all of grades 1-7; Secondary means having attended any or all of grades 8-13 or above.  
Source: De Walque (2004)

**Figure 2.2: Percentage of women who used a condom during sex in previous month**



Source: GCE graphic using DHS data from www.statcompiler.com

**Figure 2.3: Percentage of men who used a condom with a recent non-regular partner**



Source: UNAIDS/WHO graphic using DHS and UNICEF data from www.macrointernational.com

**Tests of cognitive achievement are incomplete proxies for the quality of education**

Educational Quality (SACMEQ), which since its first survey in Zimbabwe in 1991 has expanded to fifteen African countries; the Latin American Laboratory for the Assessment of Quality in Education (LLECE), which began in 1997 and covers sixteen countries; the UNESCO Monitoring Learning Achievement (MLA) project and the survey in French-speaking Africa known as the *Programme d'analyse des systèmes éducatifs de la CONFEMEN* (PASEC).

**Comparisons across countries and over time**

Tests of cognitive achievement are incomplete proxies for the quality of education. They tell nothing about values, capacities or other non-cognitive skills that are important aims of education. Moreover, if the extent of value added by schooling, even in the cognitive domain, is to be known, such tests need to be supplemented by measures of the background characteristics that learners bring to formal education. The aforementioned studies differ in the extent to which their methodologies allow for these dimensions. They also differ as to whether comparisons over time, and across countries, can be made.

It is possible to compare learning achievement scores among the countries within each study, but not among the studies themselves.<sup>17</sup> This is partly because they examine different age or education groups: PISA, for example, focuses on 15-year-olds while the others concentrate on primary school pupils. There are grounds for believing that many of the African countries included in SACMEQ have much poorer reading achievement than the IEA countries, but, in the absence of a common scale, this cannot be properly demonstrated.

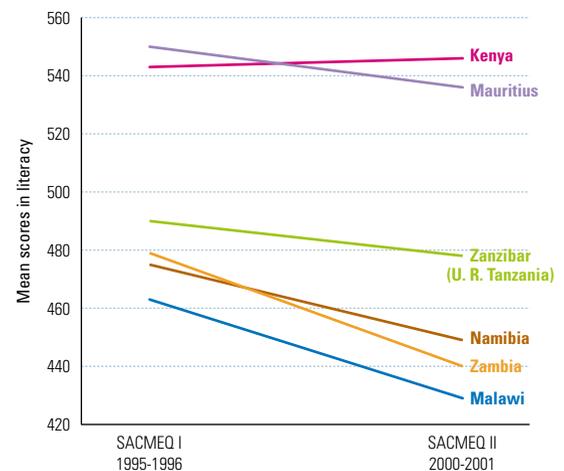
The second IEA science study was the first to allow achievement over time to be compared (Keeves and Schleicher, 1992). It found that, from 1970 to 1984, general science achievement scores at mid-secondary level increased for England, Finland, Hungary, Italy, Japan, the Netherlands, Sweden and Thailand, were roughly unchanged for Australia and decreased for the United States. The causes of the changes were unclear in the absence of information about system change and curricular modifications in these countries.

Results from PIRLS allowed comparison of changes in reading comprehension between 1991 and 2001 for the grade covering 9-year-olds. It indicated that achievement levels increased significantly in Greece, Slovenia, Iceland and Hungary, changed insignificantly in Italy, New Zealand, Singapore and the United States and fell in Sweden.

As regards developing countries, some strong and interesting comparisons emerge from SACMEQ I (1995/96) and II (2000/01).<sup>18</sup> As Figure 2.4 shows, five of the six countries included in both rounds of SACMEQ had declines in literacy achievement scores, although these differences were statistically significant only in Malawi, Namibia and Zambia. In Kenya, SACMEQ I was conducted in 1995/66–98, so the interval between the two surveys was three years rather than four to six years for the others. Here again there was no significant change. The bold line showing the average for all six countries indicates a 4% decline in achievement scores. In view of the comparatively short period covered, the consistency of these results across the region is compelling.

Table 2.2 shows changes in important contextual circumstances for the sampled schools in Malawi, Namibia and Zambia. The age of pupils in grade 6 decreased over the period, which in some circumstances could have a bearing on

**Figure 2.4: Changes in literacy scores between SACMEQ I and II in six African countries**



Source: Postlethwaite (2004)

17. Some analysts are testing such comparisons (e.g. Hanushek, 2004; Pritchett, 2004; Crouch and Fasih, 2004), but interpreting them requires making strong assumptions.

18. Measurement across time is also a major focus of the PISA research, but the results comparing the 2003 testing with that of 2000 were published too late for inclusion in this Report.

**Table 2.2: Percentage and mean differences in selected variables between SACMEQ I and II**

Variable	Malawi	Namibia	Zambia
Pupil age in months	-7.1*	-11.9*	-4.9*
Pupil sex, % female	1.3	0.7	2.6
Pupil possessions	-0.04	-0.04*	-0.07*
Parental education	0.2	0.1	0.2
% sitting places	21.4*	-2.0	5.4*
% writing places	26.0*	1.4	32.2*
Own reading book	-5.6	-5.9	0.7
Teacher age in years	1.7	1.5	4.0*
Teacher sex, % female	1.8	-8.6	13.5*
Teacher years experience	0.9	0.7	3.8*
School resources (22)	-0.42	0.10	0.15
Class resources (8)	0.7	-0.3	0.0
Lack of pupil materials	0.09	-1.05	0.30

Note: Asterisks indicate that differences were statistically significant at the 95% level of confidence.

Source: Postlethwaite (2004)

performance. Here, however, the number of over-age pupils was declining – a fact that could be counted as progress from the perspective of ministries of education and that was unlikely to lead to negative reading performance. On the other hand, household income (measured by whether pupils had particular possessions or amenities at home<sup>19</sup>) also appears to have fallen over the period, particularly in Namibia and Zambia – indicating economic decline or enrolment of pupils from poorer homes, or both.

Some aspects of the school environments appear to have improved over the period. A significantly higher proportion of pupils in Malawi and Zambia had their own seats and a desk or table on which to write. The age and experience of teachers were higher in Zambia, and a greater proportion of them were female (not the case in the other countries). The percentages of pupils having their own textbooks (i.e. not having to share) were virtually unchanged, however, as were the schools' physical resources.<sup>20</sup>

Overall, what accounts for the decrease in achievement in these three countries is not entirely clear. It is likely that the reduced average income of pupils' households was a factor in all three cases. In Zambia, per capita income declined sharply during the 1990s and demands on pupils to supplement incomes – at the cost of their school performance – probably increased. In Namibia, a higher proportion of poorer

households were sending their children to school. In Malawi, rapid expansion, which led to the number of primary pupils almost doubling over the decade, was a significant factor in the qualitative decline. The abolition of school fees there led to a much greater proportion of children from lower socio-economic backgrounds attending school. In addition, Malawi's performance on school resources was the lowest for all six countries. It was significantly worse than in Namibia and Zambia and had fallen in absolute terms over the years between SACMEQ I and II.

### More general explanations for pupil achievement

Each study made great efforts to identify the major factors influencing achievement. What were the main results? In nearly all education systems, pupils' home background was found to be important. Those from higher socio-economic backgrounds – where parents had more education and households had more material possessions, including more books – tended to perform better than those from poorer homes. In the African and Latin American studies there were also strong urban-rural differences, reflecting both higher incomes and better education facilities in urban areas.

In many developing countries, the material resources in schools are inadequate. In the SACMEQ studies the average child was in a school with 8.7 of the twenty-two desirable school resource items; the range was from 4.3 items in Malawi to 16.7 in the Seychelles, with wide urban-rural variation within countries. Even in countries that had achieved some degree of equity in the provision of material resources, the teachers in urban schools tended to be better qualified and more experienced than those in rural areas. Some schools did not even have enough seats for all pupils. On average, for all fifteen SACMEQ countries, 10% of pupils lacked a place to sit. By country the proportion ranged from 45% in Zanzibar to zero in Botswana, Lesotho, Mauritius and Seychelles.

The sex of primary teachers has an influence on performance, particularly of girls. The SACMEQ studies showed wide variations by country. For all countries, 53% of pupils surveyed were taught by female teachers, on average, but the share

**In all fifteen SACMEQ countries, 10% of pupils lacked a place to sit**

19. Pupils were asked which of the following they had at home: daily newspaper, weekly or monthly magazine, radio, TV set, video recorder, cassette player, telephone, refrigerator, car, motorcycle, bicycle, piped water, electricity (main, generator or solar) and a table at which they could write.

20. School heads were asked which of the following were available to them: school library, school hall, staff room, head's office, storeroom, cafeteria, sports area/playground, school garden, piped water/well or borehole, electricity, telephone, first-aid kit, fax machine, typewriter, duplicator, radio, tape recorder, overhead projector, television set, video recorder, photocopier and computer.

### Pupils taught by females scored higher than pupils with male teachers

ranged from 17% in Uganda to 99% in Seychelles. Pupils taught by females scored, on average, three-tenths of a standard deviation higher than pupils with male teachers.

Other items prominent in the African studies were behavioural problems of pupils (and teachers): late arrival, absenteeism and pupil dropout were all correlates of poor performance. In the PISA studies, where socio-economic advantage improved performance, changes in the school climate, teacher morale and commitment, school autonomy, teacher-pupil relations and disciplinary regime had some compensatory influence towards greater equity. In the Latin American countries covered in the LLECE studies, pupil socio-economic background and classroom climate appeared to be the most important predictors of achievement.

Box 2.2 summarizes some of the major findings from more than forty years of research

conducted through the IEA programme. Three have particular importance for policies aimed at improving education quality. First, the distribution of abilities in the population has a significant impact on average achievement levels. The greater the overall proportion of children enrolled, the lower average achievement levels tend to be. The achievement levels of particular cohorts of ability, however, are not affected – the cognitive achievement levels of the most able decile are unchanged by expansion. Second, time spent actually working on particular subjects, either in school or as homework, affects performance, especially in mathematics, science and languages. Third, although socio-economic status is influential in determining achievement in all contexts, textbook availability and school resources appear to be capable of countering socio-economic disadvantage, particularly in low-income settings.

#### Box 2.2 Major conclusions from more than forty years of international achievement surveys

Results of the IEA studies, now covering fifty countries and carried out over more than forty years, suggest the following conclusions:

- Marked differences exist between average levels of pupil achievement in the industrialized countries and those in less developed countries (LDCs) even though not all pupils in the various school-age groups were enrolled in the LDCs.
- The average level of achievement within a country at the terminal secondary school stage is inversely related to the proportion of the age group enrolled (or the age group studying the subject surveyed).\*
- At the terminal level, when equal proportions of the age group are compared, only small differences in levels of achievement are found, irrespective of the proportion of the age group enrolled at that level. Thus, the best students do not suffer as retention rates increase.
- Student achievement in mathematics, science and French as a foreign language is positively related to the time spent studying the subject at school, both across and within countries.
- Student achievement in mathematics, science and French as a foreign language is also positively associated with the time spent on homework, after other factors influencing achievement are taken into account.

- The average level of student achievement across countries is positively related to the time spent in class studying the content of the items tested.
- The impact of increased textbook use on student learning in LDCs is strong. The same effects are not detected in richer countries, probably because of the wider availability of textbooks in those countries.
- Measures of the socioeconomic status of pupils' families are positively related to student achievement in all countries, at all age levels and for all subjects.
- Although the effects of home background variables on student achievement are similar for all subject areas, the effects of learning conditions in the schools differ by subject and are sometimes equivalent to or greater than the influence of home background on student achievement.

\* Among the participating countries, the correlation between the proportion of an age group enrolled in a particular grade and the average measured achievement in mathematics and science, and that between the proportion of an age group specializing in one of the sciences and achievement in it, range from -0.69 and -0.88.

Sources: Keeves (1995: 2-23); Mullis et al. (2003: 36-38); Postlethwaite (2004).

## What determines quality? Lessons from eleven countries

Lessons to aid understanding and improvement of the quality of education can be gained in a wide range of ways. Quantitative international analysis, discussed above, is one approach. Country case studies provide a different way of profiting from national experiences, using both qualitative and quantitative information. In this section, eleven countries have been selected for such analysis. Four of them – Canada, Cuba, Finland and the Republic of Korea – have achieved high standards of education quality. They have shown leading performance in achievement surveys,<sup>21</sup> a criterion that admittedly is just one aspect of education quality. Cuba and the Republic of Korea have achieved high standards in the past two or three decades, the former inspired by a strong belief that education helps it achieve the objectives of its 1959 revolution, the latter viewing education as fundamental to its post-war economic expansion. In Finland, an industrialized country with a longstanding tradition in education (Finland Ministry for Foreign Affairs, 2002), the economic crisis of the 1990s provided a more recent impetus for a knowledge-based economic strategy.<sup>22</sup> Canada is a country where immigration has underpinned socio-economic development and education is judged to be the key to nation building.<sup>23</sup>

The seven other countries discussed – Bangladesh, Brazil, Chile, Egypt, Senegal, South Africa and Sri Lanka – have demonstrated strong commitment to EFA. All are developing countries that have been successful in expanding access to primary education. They have also made strong progress towards gender parity, or have achieved it. All seven have made big efforts to increase the quality of education in terms of learning outcomes, even if substantial, measurable progress has not yet materialized in all cases.

Senegal and Bangladesh are the poorest of these countries. There, the major challenge is to close the enrolment gap in primary education while at the same time addressing quality. In both countries the non-formal sector plays an important role. For Sri Lanka and Egypt, universal primary enrolment (UPE) is within reach; the keys to raising quality include greater consultation with civil society on national reforms (Sri Lanka) and a business approach (Egypt).

In South Africa, equity has driven education development over the last decade. Brazil and Chile, countries on the threshold of industrialization, are aiming to improve education further through large projects (Brazil) or major financial investments (Chile).

Table 2.3 and Figure 2.5 compare the educational profiles of the eleven countries. It can be seen that the four 'high-performance' countries are ahead of the others in terms of primary and secondary enrolment. Their pupil survival rates to grade 5 also tend to be higher, and the gender ratios of enrolment more equal, than in the seven countries here labelled 'ambitious'. However, when comparative performance on international achievement tests is examined, more substantial gaps appear, in general between high-performing and less developed countries.<sup>24</sup>

### The four 'high performance' countries: how did they do it?

For several decades, several South-east and East Asian countries pursued a strategy of building a larger 'stock' of trained human resources than strictly needed in the short term, in order to attract knowledge-intensive investment and thus boost economic expansion.<sup>25</sup> The Republic of Korea was one of this group. As early as 1959 it had managed to enrol 96% of its children in primary school. The following three decades saw rapid development of education, large increases in the availability of educated youth and adults, and sustained economic growth.

By 1980, the Republic of Korea had shifted its emphasis in education from expansion towards a focus on quality, giving more importance to students' 'sense of the future and ... social and moral responsibilities' (KEDI, 1979). The exploding demand for schooling had resulted in overcrowded classrooms<sup>26</sup> and excessive competition for scarce places in secondary and tertiary education. The degree of competition was felt to be harmful for learners and parents. Distance education and adult education were expanded during the 1980s, to ease the pressure on the regular school system. Entrance examinations were reformed or abolished. Teachers received longer training and better incentives, while physical facilities in schools were improved. An infrastructure of research institutes at national level – including the Korean

21. Canada, Finland and the Republic of Korea, as this chapter notes, performed well in the first round of PISA, for which data were collected in 2000 (OECD/UNESCO Institute for Statistics, 2003), and in various earlier international assessments. Cuba participated in a survey conducted by OREALC, the UNESCO regional office for Latin America and the Caribbean (Casassus et al., 2002). The average performance level of its pupils was remarkably high compared to that of other countries in the region.

22. The developments in the Soviet Union in the early 1990s caused a deep crisis in Finland: GDP dropped by 14% and unemployment rose to 20%. Yet, by 2003, its economic competitiveness had been restored. Knowledge – particularly education and research and development – is said to have played a major role in the recovery. See, for example, [www.warsawvoice.pl/view/4268](http://www.warsawvoice.pl/view/4268).

23. The children of immigrants in Canada perform better on standardized tests than those in other countries participating in PISA (Döbert, Klieme and Sroka, 2004).

24. Pritchett (2004) found, for instance, that only the 3.2% best-performing Brazilian children were better in mathematics than the average Danish pupil. Comparisons of Indonesia with France and of Peru with the United States revealed very similar differences between these middle- and high-income countries.

25. Wade (1990) and the World Bank (1993) indicate that several countries, including the Republic of Korea, Singapore, Taiwan and Hong Kong, have seen rapid economic expansion associated partly with certain monetary and macroeconomic policies, and partly with public investment in social infrastructure, particularly education (see also Republic of Korea, 2003: 12 and Corrales, 1999: 24–5.)

26. In 1978, the rural average was 44.8 pupils per class and the urban average was 61.0 (excluding Seoul and Busan, where the figure was 71.8). The national average was 52.8; for comparison, the figure in Japan was 32.9 and in China 47.9 (KEDI, 1979: 77). Reducing class size has been a priority ever since. In 2004, classes are expected to have thirty-six pupils or fewer, according to the Presidential Commission on Education and Human Resource Policy (2002: 61). The commission recommends adding assistant teachers for classrooms with more than thirty pupils.

**Table 2.3: Education system and background characteristics in eleven countries**

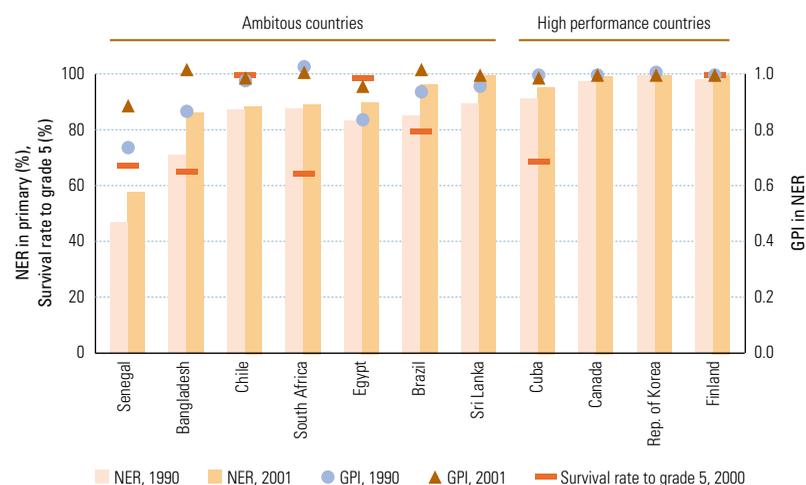
	Background		Adult literacy <sup>1</sup>		Primary		Secondary		Teaching staff		Finance
	Population in thousands	GNP per capita (PPP)	Total %	GPI (F/M)	GER (%)	GPI (F/M)	GER (%)	GPI (F/M)	Primary school teachers, female %	Pupil/teacher ratio in primary	Total public expenditure on education as a % of GNP
<b>Ambitious</b>											
Bangladesh	140 880	1 710	41.1	0.62	97.5	1.02	46.9	1.10	36.0	55.1	2.2
Brazil	174 029	7 350	86.4	1.00	148.5	0.94	107.5	1.10	92.1	23.0	4.2
Chile	15 419	9 240	95.7	1.00	102.7	0.98	85.5	1.02	77.6	32.2	4.0
Egypt	69 124	3 720	55.6	0.65	96.9	0.94	88.1	0.93	53.3	22.5	...
Senegal	9 621	1 540	39.3	0.61	75.3	0.91	18.7	0.67	22.8	48.9	3.2
South Africa	44 416	9 530	86.0	0.98	105.1	0.96	86.4	1.09	77.8	37.1	5.8
Sri Lanka	18 752	3 380	92.1	0.95	110.4	0.99	80.8	...	...	...	1.3
<b>High performance</b>											
Canada	31 025	28 570	...	...	99.6	1.00	106.2	0.99	68.1	17.4	5.3
Cuba	11 238	...	96.9	1.00	100.3	0.96	89.1	0.99	78.9	13.5	8.7
Finland	5 188	25 500	99.7 <sup>2</sup>	0.99	102.0	0.99	126.5	1.11	74.2	15.6	6.4
Republic of Korea	47 142	15 060	98.0	0.98	102.1	1.00	91.1	1.00	71.6	32.0	3.6

Notes: see source tables for detailed country notes.

1. Adult literacy rate is for the period of 2000-2004.

2. Adult literacy rate for Finland is based on the rates of 0-3 years of schooling.

Source: Statistical annex, Tables 1, 2, 5, 7, 8 and 14. Adult literacy rate for Finland: European Social Survey ([www.europeansurvey.org](http://www.europeansurvey.org)).

**Figure 2.5: Performance indicators for primary education in eleven countries**

Note: For Cuba, the survival rate is to the *last grade* (not grade 5, which is unavailable).

See source tables for detailed country notes.

Source: Statistical annex, Tables 5 and 7.

Educational Development Institute – served to guide this reform process, while an education tax was introduced to finance it. Through the 1990s, these initiatives were consolidated – a process reinforced by the founding of advisory bodies that transcended political regimes and sought consistency in education policy (Republic of Korea, 2003: 23–5).<sup>27</sup>

Notwithstanding these efforts, class size remains large, even though it has been almost halved. In an important sense, being taught in a smaller group is a quality gain in its own right. Whatever disadvantage larger classes entail, though, it is overcome in the Republic of Korea through the willingness of pupils (and parents) to make extra efforts and through pedagogies and classroom climates that facilitate learning in large groups. Still, the Republic of Korea's first place for science, third place for mathematics and seventh place for reading in the PISA study covering forty-one high- and middle-income countries (OECD/UNESCO Institute for Statistics, 2003) are remarkable, especially given PISA's adoption of more contextual (less 'schoolish') testing methods. These rankings suggest that the Republic of Korea has reached a broader interpretation of learning outcomes than other countries, including many at higher income levels.

At an earlier stage than the Republic of Korea, Cuba was emphasizing education's role in developing the whole individual (including physical education, sports, recreation and artistic education) while explicitly linking education with life, work and production (Amadio et al., 2004). Following the Cuban Revolution, education and health care were strongly prioritized (Ritzen,

27. See also Corrales (1999: 27) and Amadio et al. (2004).

1999) to support human development. They were seen both as desirable ends in themselves and as a means of assuring the country's economic and political independence. The balance between these two types of value, which the Republic of Korea sought in the 1980s, was present in Cuba from the start of its development process.

In both countries, competition plays an important role, but it does so in very different ways. While competition for places arose as an unintended effect of scarce provision in the Republic of Korea, educational opportunities in Cuba have been abundant and freely accessible at each level, partly because investment in education is high, amounting to 10–11% of GDP (Gasperini, 2000: 7). The Cuban form of competition, 'emulation', is conceived of as self-improvement through solidarity and collaboration among peers. Emulation occurs among pupils, among teachers and among schools. For each group, incentives reward excellence, and mechanisms are in place to make sure others benefit from the experience. An example is the *colectivo pedagógico*, a group of subject teachers meeting frequently for mutual learning and joint development of curricula, methods and materials (Gasperini, 2000: 9–14). The result is an education system that stakeholders are encouraged to improve. Extra-curricular contributions, such as to school maintenance, are common, and the system is characterized by a high level of discipline and classroom order (Carnoy, Gove and Marshall, forthcoming). Cuba's educational feats are impressive: it reduced illiteracy from 40% to near zero in ten years (Ritzen, 1999), and in the recent OREALC/UNESCO study the average performance of the bottom quartile of its tested students was higher than the average of the top quartile for any other country in the survey.<sup>28</sup>

Can other countries emulate Cuba's policies? The revolutionary spirit that inspired teachers, pupils and parents to make great efforts for the benefit of the schools may prove to be unique. Nevertheless, learning communities of teachers can be created, and schools' performance scrutinized, in other contexts.<sup>29</sup> The high esteem in which the teaching profession is held in Cuba seems crucial to its success. This could be a potentially vulnerable point: as the country opens up to tourism and foreign investment, teacher salaries – very low when expressed in foreign currency – might not be able to compete with

salaries in other sectors (Gasperini, 2000: 21). Moreover, the availability of relatively well-paid tourism jobs for which few qualifications are needed could encourage youth not to continue their education.

Canada<sup>30</sup> is another country where the teaching profession is held in high esteem. Despite teacher shortages, admission to teacher training is highly competitive, and only 10% of applicants succeed. Even pre-primary teachers need a university degree. In-service training amounts to forty days per year in some parts of the country; participation is often obligatory or a condition for promotion, and is financially rewarded. A system of accreditation in Ontario, which is under consideration elsewhere, tests teachers every five years, and those who fail lose their teaching certificate.

Canada not only maintains high standards for teachers, it also provides a well-developed support system. Its school development teams (at district level) and school advisory councils (bringing local stakeholders together in support of the school) are reminiscent of some Cuban institutions. Monitoring is a third hallmark of Canadian education. A culture of indicator use has developed at all levels. The performance of students, schools, districts and even provinces is tracked closely. This is seen as a way of revealing both excellence and underperformance and as a basis for designing policy interventions.

Despite a decline in investment in education from 9% of GNP in 1970 (Wisenthal, 1983) to 6.6% of GDP in 1999, Canada performed very well in PISA, ranking second for reading, seventh for mathematics and sixth for science (out of forty-one high- and middle-income countries). Most notable is the excellent performance of immigrant children in Canada compared to those in other industrialized countries. This underlines one of the characteristic objectives of Canadian education: to build a nation while cherishing the cultural diversity of the population.

Finland<sup>31</sup> had the highest overall outcomes in the PISA tests, ranking first out of the forty-one countries for reading, fifth for mathematics and fourth for science. Disparities in student achievement are very small, as is the impact of social background on achievement (Väljäärvi et al., 2002: 28). That is what Finland intended

**The high esteem in which the teaching profession is held in Cuba seems crucial to its success**

28. We should continue to bear in mind the point made in footnote 1, however: that achievement scores are far from the whole picture.

29. See, for example, the discussion of Canada below, and UNESCO (2003a: 216).

30. Unless indicated otherwise, the information on Canada is derived from a comparative study of seven countries, based on PISA, written by several national teams and edited by Döbert, Klieme and Sroka (2004).

31. Except where indicated otherwise, the information on Finland is also derived from Döbert, Klieme and Sroka (2004).

**The combination of high performance and moderate expenditure has made Finnish education an interesting benchmark for many countries**

in investing consistently and for many decades in human development, pursuing both equality of opportunity and inclusion. Economic competitiveness and education sector performance, however, are also key areas in the country's education strategy for 2015 (Finland Ministry of Education, 2003). This emphasis on the more utilitarian objectives of education was absent in the 1980s (Amadio et al., 2004) and may be a consequence of Finland's economic crisis in the 1990s after the collapse of the Soviet Union. Finland consciously chose a knowledge-based recovery strategy, but could not afford high investment in education. The 5.8% of GDP that it invests in education is only slightly above the OECD average and clearly below Scandinavian standards. The combination of high performance and moderate expenditure has made Finnish education an interesting benchmark for many countries.

As in Canada, selection for teacher training is very rigorous. Every teacher has masters' degrees in two subjects. Other factors that are said to explain Finland's high performance in PISA are its comprehensive pedagogies, students' own interests and leisure activities, the structure of the education system, school practices and Finnish culture<sup>32</sup> (Väljäävi et al., 2003: 4).

The experiences of these four countries suggest three common characteristics. The first concerns teachers. High esteem for the teaching profession, thorough pre-service training and sometimes restrictive admission, and a well-developed constellation of in-service training, plus mechanisms for mutual learning and teacher support are evident in all these countries. There are no concessions on teacher quality, even where teacher shortages exist.

The second is continuity of policy. The Republic of Korea consciously sought to neutralize the impact of political change by establishing advisory bodies. In Cuba, continuity is implied in the political system. Canada and Finland have strong education knowledge bases (within institutions for teacher training and support) that seem to prevent governments changing course too frequently and radically.

The third characteristic is the high level of public commitment to education, which seems to emanate from a strong political vision. The

Republic of Korea's determination to become and remain globally competitive, Cuba's will to defend the revolution, Canada's belief that its strength as a nation lies in cultural diversity, and Finland's deep commitment to human development and equality – each, in its own way, has profoundly affected education policies and outcomes.

One other characteristic, in the Republic of Korea and Cuba, is an extremely high level of energy among learners, teachers and parents. In both countries it is associated with an atmosphere of competition, albeit from very different standpoints and in very different forms. Whether and how this can be mirrored in other developing country contexts is an open question.

**The seven ambitious countries: what can they learn?**

Senegal<sup>33</sup> has been strongly committed to basic education and has rapidly expanded access. Between 1990 and 2000, its net enrolment ratio rose from 48.2 to 63.1, with the gender parity index increasing from 0.75 to 0.90. The country is now looking for a better balance between quantity and quality. Indicators of quality still seem to lag behind, however, with relatively high repetition rates in the higher grades, a low ranking in the PASEC survey<sup>34</sup> and little progress according to the national assessment system – the *Système national d'évaluation des rendements scolaires* (SNERS). Transition to secondary education is low compared to other countries in sub-Saharan Africa. Unlike the Republic of Korea, Senegal has not been able to benefit from economic growth when raising quality; in fact, growth was negative between 1990 and 2000.

With limited money to fight illiteracy and with current efforts to do so widely scattered, the government adopted an innovative approach called *faire-faire*, or 'getting it done' (Niane, 2004: 12). It is based on bringing relevant partners together, sharing out duties and responsibilities to local actors, giving a voice to stakeholders and decentralizing the education system. The *faire-faire* approach has been applied to literacy programmes in which over a million adult women were enrolled in 2003, along with almost half a million boys and girls who missed out on regular school. While achievement in regular education was stagnant, progress in

32. The importance of Finnish culture and students' interests and leisure activities points to the cumulative intergenerational effects of sustained support to education. It is known that student performance is driven by the cultural capital and the interests students bring from home to the classroom, which in turn are strongly linked to parent's education attainment. Thus, if a country invests long enough in good education, its quality may to some extent become self-sustaining.

33. Unless otherwise indicated, the information on Senegal is from Niane (2004) and the Statistical annex of this Report.

34. PASEC, discussed in the section on international assessments, was conducted in Burkina Faso, Cameroon, Ivory Coast, Madagascar and Senegal in 1996 and (for Madagascar) 1998. The results for Senegal were the lowest among the five countries.

these literacy programmes was remarkable between 1998 and 2001. Another innovation is the formation of *collectifs des directeurs*, or regional groups of school principals, which echo the Cuban concept of stimulating mutual learning among peers and provide a way to make decentralization work. Senegal realized that school autonomy should not leave schools in isolation and that networks are important.<sup>35</sup>

On balance, the Senegalese experience seems to reflect a sense of dissatisfaction with the bureaucratic nature of the formal school system. Schools remain passive, expecting initiatives to come from above rather than be self-generated. Hence, there is a tendency to operate outside the formal system (*faire-faire*) or independently from it (*collectifs des directeurs*) to accomplish change. While some high-performing countries pursue similar strategies, few of them made this choice at such an early stage of education development. The chances for a robust, vision-led, consistent national education strategy thus seem limited, although the Education and Training Programme (2000–10) does chart a course for the longer term. One strong point is the high proportion of teachers in Senegal with good academic backgrounds. The number of teachers who actually have a teaching certificate, however, has fallen.

Bangladesh<sup>36</sup> also made impressive progress in access to primary education between 1990 and 2000, possibly as a result of the advent of democratic government at the beginning of that decade. In this large and poor country – it has 130 million inhabitants, half of whom live below the poverty line – the net enrolment ratio went up from 71.1 to 88.9 over the decade and gender parity was achieved by the turn of the century. The ten-year rise in the gender parity index from 0.87 to 1.02 implies an increase in quality in its own right. But, as in Senegal, learning achievement is still a problem. Although it is not possible to assess developments over time,<sup>37</sup> the proportion of pupils demonstrating specified minimum levels of mastery for basic subjects may be as high as 50%, as reported by the government (Latif, 2004: 9), or as low as the 9% that independent sources report (Education Watch, 2000). That three-quarters of pupils nevertheless complete primary education points to weaknesses in curricula, teaching and quality assurance.

Another similarity with Senegal is the large non-formal education sector. NGO-supported schools in Bangladesh cater for 2 million children, against 19 million in the formal system. BRAC (formerly the Bangladesh Rural Advancement Committee) is an NGO focusing on children who did not enrol at the normal entry age. Teachers receive very short preparatory training (twelve days) but ongoing support is well-organized and supervision is given weekly. BRAC children perform significantly better on life skills and writing than their peers in normal schools and do equally well in reading and numeracy. In 1999, transition to secondary school was also higher for BRAC children: 95.3%, against 81.9% for children in the formal system.

The non-formal sector is large and diverse. With eleven types of primary education and many partners, it is hard to develop and implement a robust national policy, the more so because the education data system does not yet allow close monitoring and precise target setting. The low student attendance rate (62%), high teacher absenteeism and relatively low number of contact hours (World Bank/Asian Development Bank, 2003), together with relatively low interest in teacher training (Latif, 2004: 8), stand in contrast to the high levels of engagement in Cuban schools and those of the Republic of Korea. Nonetheless, the government of Bangladesh is continuing its strong commitment to education reform into the second phase of the Primary Education Development Programme.<sup>38</sup>

Sri Lanka<sup>39</sup> is making progress in both access and quality. UPE is within reach, and gender parity has been achieved. Grade repetition and dropout have declined rapidly and the promotion rate stood at 98.4% in 2001. The pupil/teacher ratio (PTR) fell from 24:1 to 22:1 between 1992 and 1999. Exclusion of Tamil children in the country's tea plantation area is being addressed. Expenditure as a percentage of GDP has been constant in recent years at around 3%, but education's share in the government's budget has fallen, in line with a decrease in the school population. Indications of achievement diverge strongly, as in Bangladesh. While evaluation of a pilot project in the Gamphala district found that 80% of the students reached mastery level (Little, 2000), a national study found 37.2% at mastery level for literacy and 22.6% for numeracy (National Institute of Education of Sri Lanka, 2002).

**Sri Lanka is making progress in both access and quality. UPE is within reach, and gender parity has been achieved**

35. See also UNESCO (2003a: 216).

36. Unless otherwise indicated, the information on Bangladesh is derived from Latif (2004) and the Statistical annex.

37. Several assessments of learning achievement have taken place in Bangladesh in recent years, but as a result of methodological differences it is not possible to compare these and identify trends. In fact, the outcomes of the various studies are remarkably different; see Latif (2004: 9–10).

38. The first phase, from 1997 to 2003, comprised twenty-seven projects. Its evaluation informed the design of the second phase (2003–08). A pool fund established jointly by donors and the government will support implementation of phase II.

39. Unless otherwise indicated, the information on Sri Lanka is derived from Peiris (2004) and the Statistical annex.

**For South Africa  
the challenge has  
been to address  
access, equity and  
quality at the  
same time**

However, recent years have seen a modest upward trend in achievement and in examination results, especially among poor children, although less so among the Tamil population in conflict-ridden northern Sri Lanka.

There are good prospects for further advances. Achievement levels for 2002 may not yet fully reflect the primary education reforms instituted in 1998 in response to alarm over Sri Lanka's international standing resulting from 1997 primary education test scores. The reform process involved consultation and mobilization of society (including children). The comprehensive approach addressed all issues believed to make a difference for quality and focused on changing teacher behaviour and developing built-in monitoring and evaluation capacity.<sup>40</sup>

Egypt's stage of educational development is comparable to that of Sri Lanka. Expansion of access between 1990 and 2000, in the largest education system of the Arab world, brought UPE and gender parity within reach (Arab Republic of Egypt Ministry of Education, 2002). Quality was placed high on the agenda. While grade repetition and dropout have been declining since the early 1990s, achievement tests did not indicate progress between 1997 and 2001 (World Bank, 2002a). But performance may soon show the effects of measures taken in the late 1990s, such as increased expenditure on teacher salaries, modest reductions in class size, 'aggressive' in-service teacher training and strengthened support systems. The latest five-year plan (2002–07) involves further efforts to achieve 'excellence for all' in 'beautiful, clean, developed and productive schools' as the government puts it (UNESCO, 2003c).

Egypt is remarkable for the business-style approach to quality assurance that it has chosen.<sup>41</sup> Possibly inspired by foreign donors, the country sees the challenge of improving education as one in which clear targets need to be set (Arab Republic of Egypt Ministry of Education, 2002: 128–134). Decentralized management based on information and communications technology (ICT) is expected to be key in meeting these targets. 'School improvement plans' unite all elements of the strategy. Periodic evaluation of school performance and participation by civil society are seen as instrumental in holding schools

accountable. With its businesslike approach, Egypt seems to have a strong vision of educational development. This vision may stabilize education policy over the longer term, as was the case for the four high-performance countries discussed above. One area in which Egypt stands out is that of its emphasis on early childhood care and education (ECCE). It has set an intermediate target of enrolling 60% of 4- to 6-year-olds by 2010, and intends kindergarten to eventually be an integral part of the system of free compulsory education.

For South Africa<sup>42</sup> the challenge has been to address access, equity and quality at the same time. The 1990s saw the net enrolment ratio stagnant at just below 90%, while the gender parity index dropped slightly, from 1.02 in 1990 to 0.98 in 2000. Apartheid left the country with large inequalities along racial lines, reflecting stark differences between rich and poor schools and districts. Results from the SACMEQ, MLA and TIMSS surveys revealed striking underperformance of South African youth compared to other countries in the region – even those investing less in education. South Africa's own expenditure decreased from a relatively high 7.3% of GDP in 1991/92 to 5.5% in 2001–04.

Measures taken soon after the abolition of apartheid in 1994 included ending resource allocation based on racial criteria and democratizing control over schools by introducing school governing bodies. In addition, a remarkable teacher redistribution and deployment project aimed at moving the better teachers to the poorer schools. Policy makers had to arrive at a new balance between public and private resources. Nine years of basic education was made compulsory, and expenditure on schooling was equalized for all racial groups. To avert flight to private schools, the government introduced an approach to financing whereby schools could introduce fees to supplement state resources, if their governing bodies deemed it appropriate. No child, however, can be denied access to schooling on grounds of inability to pay the fees, and all or part of the amount is waived for children from families whose income is less than thirty times the fees charged. As intended, this policy has kept fees from deterring enrolment among the poor, while also maintaining participation by richer families in the state education system (Fiske and Ladd, 2004).

40. For a detailed account and evaluation of the education reform process in Sri Lanka, see Little (2003o).

41. See also the Hungarian experience (UNESCO, 2003a: 216).

42. Unless otherwise indicated, the information on South Africa is from Chisholm (2004) and the Statistical annex.

A new curriculum was launched in 1997. To accelerate its use and enhance its impact on quality, a plan for 2000–04 gave special attention to building teacher self-esteem, developing a professional body for teachers and providing better support. The central role accorded to securing the commitment and quality of teachers is consistent with the lessons from the high-performance countries discussed earlier.

Brazil<sup>43</sup> and Chile<sup>44</sup> are among the countries that symbolize the enormous progress towards EFA in Latin America and the Caribbean. While Cuba is clearly the region's star, Brazil and Chile, along with Argentina, performed relatively well in the OREALC/UNESCO survey.<sup>45</sup> But, measured over time, achievement has been stagnant in Brazil and Chile, despite ambitious efforts to improve.

In the case of Brazil this must be seen against the backdrop of expansion: NER rose from 86.4 in 1990 to 96.7 in 2000. Most new entrants are likely to come from marginalized groups. Their 'school-readiness'<sup>46</sup> will generally be lower than that of children from groups that have been included in education longer. Keeping average achievement constant over the period of expansion can thus be seen as improvement. Moreover, the achievement gap between the highest and lowest performers is relatively small in Brazil. The country has made strong policy efforts to address regional and social inequalities in inputs (especially funding) and achievement. Two flagship projects under the Ten-Year Plan for Education For All (1993–2003) were FUNDEF and Bolsa Escola/PETI. The former is an equalization fund that reduced regional inequalities in the funding of primary education, despite the complex three-level governance structure (local, state, federal). Funding for the poorest schools increased significantly, explaining the growth of enrolment and boosting teachers' salaries and training. Bolsa Escola and PETI are widely praised initiatives to boost schooling among poor families (UNESCO, 2003a: 169). Other projects under the aegis of the ten-year plan are FUNDESCOLA, to enhance school attendance and combat dropping out; PROFORMACAO, to train unlicensed teachers through distance learning; the National Schoolbook Programme, which increased the number of textbooks in terms both of copies and titles; School Cycles implying automatic promotion from first to second grade; and

Accelerated Learning Programmes, to give late entrants a chance to catch up. These illustrate a difference between the ways Brazil and South Africa address inequalities. Brazil has selected several large projects, each addressing an aspect of the education system or its context. South Africa has followed a broader strategy, focusing on equalizing resource inputs among schools, and making teacher quality a key element of reform.

Chile, like Finland, consciously chose education as a core strategy for socio-economic development. In the early 1990s, the former military regime had left the country with the challenge of revitalizing and modernizing the economy, enhancing social cohesion and building a democratic citizenry. Public expenditure on education rose from 2.4% of GDP in 1990 to 4.4% in 2001. In absolute terms it more than tripled between 1990 and 2003, and private expenditure also increased strongly. An expansion of the annual number of school hours, from 880 to 1,200, is on course, the idea being eventually to end the practice of multi-shift teaching. Social assistance has been improved, enhancing enrolment of poorer children. ICT is seen as key to modernizing education. Authoritarian teaching styles, mainly aimed at keeping order in the classroom (Carnoy, Gove and Marshall, forthcoming; OECD, 2004d: 36–7), are slowly giving way to more advanced approaches aimed at activating students. However, all this remains weakly associated with learning targets (OECD, 2004d: 36–7). Indeed, Chile's national student monitoring system (*Sistema de medición de la calidad de la educación*, or SIMCE) reveals only modest gains. Much more progress in achievement is reported from compensatory programmes such as P.900, which targets the schools in the lowest-performing decile, and the rural programme of MECE,<sup>47</sup> for small multigrade schools. Thus, as in Senegal, progress is found in special settings and has not yet spread throughout the system. Low self-confidence among Chilean teachers, reported in 1999,<sup>48</sup> may play a role in the lack of progress, given the importance of teacher esteem in the four high-performance countries.

Taking stock of the experience of the seven ambitious countries, we can observe a number of common characteristics. All seven have made significant progress in expanding access, and are

### Brazil and Chile are among the countries that symbolize the enormous progress towards EFA in Latin America and the Caribbean

43. Unless otherwise indicated, the information on Brazil is from Gusso (2004) and the Statistical annex.

44. Unless otherwise indicated, the information on Chile is from Cox (2004) and the Statistical annex.

45. Again, though, it is worth recalling that achievement scores do not give the whole picture.

46. See 'Start with learners' in Chapter 4. The extent to which children are sufficiently prepared for school varies significantly, with social background playing an important role.

47. 'Mejoramiento de la calidad y equidad de la educación' (improvement of quality and equity in education) is a series of programmes aimed at improving education at various levels.

48. This was among the findings of the 1999 TIMSS survey. Of the thirty-eight countries covered, Chile and the Philippines shared the lowest position concerning the number of teachers who are confident about their mastery of mathematics, and Chile and Italy shared the lowest position regarding science mastery. Self-confidence regarding teaching skills was also lower in Chile than in most of the other countries.

**Successful qualitative reforms require a very strong leading role by government and a robust long-term vision for education**

now shifting the balance between quantity and quality. Indeed, raising quality generally seems to *follow* the expansion of access, up to a certain level. While the achievement of higher quality is implied by progress on gender parity and better-equipped schools, the picture as regards learning achievement is less favourable. Progress in cognitive outcomes has been modest, absent or restricted to certain projects, though some countries can expect broader, more substantial improvement as a result of recent policy initiatives.

Judging from the experience of the four high-performance countries, successful qualitative reforms, in the context of systems characterized by universal access, seem to require a very strong leading role by government and a robust long-term vision for education. These characteristics are less prominent in the seven ambitious cases, where another key requirement – having sufficiently motivated, well-supported teachers – also seems insufficiently fulfilled.

Although these characteristics and similarities are important, it seems clear from this discussion of eleven very different countries that a search for any general theory of successful educational reform is unlikely to succeed. Contextual differences limit the transferability of policy lessons from one country to another (OECD, 2004a), even among relatively comparable countries.<sup>49</sup> Usually, reforms that aim to raise the quality of education are politically even more difficult to pursue than those that aim at expansion.<sup>50</sup> This means the political context is likely to have a strong impact on prospects for reform, irrespective of the technical and resource context of the school system. The time factor also complicates the analysis: the impact of an education policy may not become apparent until years after it is implemented, and is never in isolation from other policies and trends. This makes one-to-one relationships between measures and effects difficult, though not impossible, to establish.<sup>51</sup>

## The quality of ECCE and literacy programmes

As the above review of country experience illustrates, most discussions and evidence of the quality of education focus on the school system. Very many youths and adults in developing

countries have had little or no access to schooling, yet the attention and resources given to literacy and skills development programmes are much lower than what is given to schooling. When such programmes do receive emphasis, it tends to be on the expansion of provision rather than on quality.

This section takes a closer look at the quality of the learning that takes place before and after primary school age.<sup>52</sup> It asks several questions: What can be said about the quality of programmes? What are characteristics of good programmes? How can quality be enhanced affordably? Do better-quality programmes have a correspondingly stronger impact on the individual, community and society?

### Does the quality of ECCE programmes make a difference?<sup>53</sup>

Assessing the quality of provision in early childhood care and education is more challenging than for schooling. Achievement tests, examinations and diplomas are largely absent at this level. National data showing provision and inputs are limited and often not easily comparable, although at programme level various quality assessment instruments have been developed.<sup>54</sup>

ECCE provides a good example of programmes where relative interpretations of quality are necessary. For example, in the early years of learning, parents involve themselves more intensively and in different ways than they do later in their children's education. Young children have a right to spend their first years in a peaceful, safe and playful environment. As the quality of ECCE depends strongly on programme context, it can be argued that the definition of quality in this area should vary and be subject to negotiation among parents, practitioners and policy makers.

Common sense suggests that the early years – when the brain matures, when we first learn to walk and talk, when self-control begins, when the first social relationships are formed – must be regarded as important. It also suggests that children whose basic health, nutritional and psycho-social needs are being met will develop and perform better than those less fortunate. Common sense also suggests that a child who

49. Germany is making an encouraging attempt to overcome this difficulty. Concerned about its PISA results, it conducted an in-depth study of a few relatively comparable but better-performing countries (Döbert, Klieme and Sroka, 2004). The design of this study could serve as a model for those wishing to learn from other countries.

50. See Corrales (1999) for a useful analysis of why this is so.

51. See the section below on the experimental evaluation of education policy interventions.

52. Previous Global Monitoring Reports have focused on the impact of ECCE and learning programmes, on the importance of ECCE for gender equality, and on conceptual issues (UNESCO, 2002a: 38–43 and 56–67; UNESCO, 2003a: 34–43, 84–95 and 181–8).

53. Unless otherwise indicated, the information on ECCE is from Myers (2004).

54. Examples are the Early Childhood Environmental Rating Scale (ECERS), the High/Scope Program Quality Assessment (PQA) and an observational instrument developed for the IEA Preschool Project.

develops well physically, mentally, socially and emotionally during the early years is more likely to be a happy and productive member of society than one who does not.

Research on early childhood development confirms common sense. The literature is vast and varied, encompassing research carried out by psychologists, medical doctors, anthropologists, neurobiologists, educators, sociologists, nutritionists and others. In general, this research supports the following propositions:<sup>55</sup>

- The early years of life are a key period for the development of intelligence, personality and behaviour.
- Early childhood learning and development can be enhanced by ECCE programmes.
- The effects of such programmes are likely to be greater for children from disadvantaged backgrounds than for their more privileged peers.
- Good programmes are sensitive to differences in cultural, social and economic contexts.

Given that ECCE matters and that ECCE programmes differ enormously in design and quality,<sup>56</sup> the question becomes whether their design and quality make a difference. The answer seems to be that they do, for most desired outcomes, even when the tendency for privileged families to choose higher-quality programmes is taken into account.

While most of the research compares children who followed a certain programme with those who did not, another kind of study has emerged over the last two decades that focuses explicitly on quality, comparing outcomes for children in ECCE centres that differ in the level of quality attributed to them.

A Turkish study, for example, compared children who received no form of care, those who were looked after by child minders and those who attended some type of preschool centre (Kagitçbasi, 1996). Although quality was not defined explicitly, the results were better for the third group. Another feature of this study was the inclusion of a parental education and support

component. This was found to produce important results regarding children's cognitive development and school performance as well as child-rearing practices in the family – the latter related in part to changes in the self-image and knowledge of the participating mothers.

In Mauritius, eighty-three children were assigned to a good preschool from ages 3 to 5, and matched on temperament, nutritional, cognitive, autonomic and demographic variables with a control group of 355 children given no special schooling. By age 10, the children who had attended the quality preschool showed better social skills and more organized thinking and had more friends than those in the control group. At age 17 and age 23, the positive effects were still pronounced: the researchers found these young adults to be more socially adjusted, calmer, better able to get along with peers and up to 52% less likely to commit a crime (Raine et al., 2003).

A fifteen-year IEA study in ten high- and middle-income countries<sup>57</sup> found similar results (Weikart, Olmsted and Montie, 2003). It sought to determine how process and structural characteristics of community pre-primary settings affected language and cognitive development. Based on a study of more than 5,000 children aged 4 and 7, in 1,800 different settings, it was found that language performance at age 7 was better the more autonomy children had been given in preschool and the higher the preschool teacher's educational level. Another study found that cognitive performance improves the less time children spend in whole-group activities and the more and better equipment and materials to which they have access. (High/Scope Educational Research Foundation, 2004).

Although most of these studies did not treat quality explicitly, it was nevertheless evident that even where pre-primary programmes operated with modest resources (using paraprofessionals and sometimes with unfavourable class sizes) they often showed a positive impact on children,<sup>58</sup> and work with parents seemed to be a factor in this effect. This is not to say that better quality would not have produced even better results, but it serves to emphasize that it is not always necessary to apply high and probably unaffordable quality standards uniformly in all settings.

**Even where pre-primary programmes operated with modest resources they often showed a positive impact on children**

55. See, for instance, Hunt (1961), Vygotsky (1962), Bloom (1964) and Piaget and Inhelder (1969), and more recent work on brain development (summarized in Mustard, 2002), the roots of antisocial behaviour (Rutter, Giller and Hagell, 1998), the prevention of intellectual disabilities (Ramey and Ramey, 1998), resilience and 'positive deviance' (Werner and Smith, 1982; Zeitlin, Ghassemi and Mansour, 1990) and nutrition and cognitive development (McKay et al., 1978), recent reviews by the National Research Council (2001), the Carnegie Corporation of New York (1994) and the Centre of Excellence for Early Childhood Development (2004), and volumes edited by Keating and Hertzman (1999) and Young (2002), among many others.

56. They may involve direct attention to children, or indirect attention via work with their parents, or be child-centred community programmes, or a combination of these. They may involve health, nutrition or education components, or a combination. They may be publicly or privately run. A range of curricula can be found.

57. Finland, Greece, Hong Kong, Indonesia, Ireland, Italy, Poland, Spain, Thailand and the United States.

58. Martínez and Myers (2003) found in Mexico that the larger urban classes reached better outcomes than smaller rural classes, since other factors appeared to be stronger, such as teacher preparation, resource availability, multigrade classes and the quality of management.

**The research evidence indicates that better child care for children of preschool age is associated with better cognitive and social development**

Most of the present 'success stories' of ECCE come from high- and middle-income countries where resources are less scarce. India and Nepal are exceptions. Both have low-cost programmes that nevertheless have a relatively strong positive impact on children and their families (UNESCO, 2003a: 182–3). In the Nepalese 'Entry Point Programme', a well-designed four-day training programme and a toolkit enabled mothers to educate each others' children one day a week, freeing them to generate income on the other days. In countries where achieving UPE is the main challenge, such low-cost measures offer an interesting option.

In conclusion, the research evidence indicates that better child care for children of preschool age is associated with better cognitive and social development. Organized preschool care and education, with some material resources and qualified teachers giving children stimulation and some choice of activities, seems to lead to better cognitive and social development later in life than does an absence of such programmes.<sup>59</sup> The impact of quality in ECCE appears to be important for children from all backgrounds, but particularly for the least advantaged.

### Literacy and skills development programmes for youth and adults<sup>60</sup>

The debate on the quality of literacy and skills development<sup>61</sup> has its own history, distinct from that on schooling. The importance of strong, widely distributed reading, writing and calculation skills in societies aiming for democracy, industrialization and farm modernization has long been recognized.<sup>62</sup> However, in the praxis of adult learning, many paradigms have come and gone over the years.

Until the 1950s, literacy skills were widely assumed to be general. Reading a manual was judged no different from reading a newspaper. An effective text or curriculum would thus serve all learners equally well, it was thought. This 'one size fits all' approach underlay government and other agency efforts, whether in campaigns,<sup>63</sup> national programmes or missionary classes. Although good evaluations of these early programmes are scarce, it is clear that their effectiveness varied widely.

Doubts about their efficacy prompted UNESCO to propose a new strategy in the 1960s. It was based upon the premise that literacy skills, being means to ends, required clear purposes and almost immediate applicability. The new approach was to be 'functional, selective, intensive and work-oriented'. Instead of one curriculum to fit all, many were developed to fit particular groups or occupations where literacy skills could raise productivity. The occupations ranged across household, agricultural and industrial situations in rural and urban areas, and addressed the interests of both women and men. However, a 1976 UNESCO/UNDP evaluation of these 'functional literacy' approaches found only modest benefits,<sup>64</sup> and confidence in such programmes' worth declined as a result.

Nevertheless, the idea of 'functional' literacy has remained central to subsequent approaches, and its main assumptions have been validated. Functional content has usually included rudimentary information on health, hygiene, nutrition, child care, agriculture, environmental concerns, savings, credit and other topics judged important and useful for unschooled and poor people, especially women. Many governments and international agencies have supported programmes with these central themes.

On a parallel though much smaller scale, some work has aimed at using literacy to address more political objectives. Where earlier literacy programmes sometimes espoused political content to encourage nation building, the Brazilian educator Paulo Freire sought to use literacy to generate political and social change from below. He developed a pedagogical strategy that would lead people to reflect on their predicaments and their causes. Learners were expected to use literacy to articulate their concerns and initiate political action to ameliorate their conditions. Although the strategy appealed to adult educators and to government departments of adult education, the scope for using it for political ends was limited, and its legacy was absorbed within the functional approach.

Since the early 1990s, however, ActionAid, an international NGO, has taken Freire's strategy a step further. Whereas Freire derived conventional texts and exercises from words and phrases of special significance in local

59. See Peisner-Feinberg (2004) for a recent review.

60. Unless otherwise indicated, the information on Literacy Skills Development is from Oxenham (2004).

61. Literacy and various other kinds of skill often go hand in hand, both in acquisition and application, though it remains important to distinguish them clearly (see UNESCO, 2003a: 84–6).

62. There is also increasing agreement on the rights of adults to empowering and relevant basic education and on the importance of participating in dynamic, rich, inclusive learning environments (UIE, 2004: 1).

63. See Arnone and Graff (1987) for evaluations of several campaigns.

64. For example, in the United Republic of Tanzania out of the 466,000 people enrolled in a literacy project carried out between 1967 and 1973 within the framework of the Experimental World Literacy Programme, 293,000 took the final examination and 96,000 passed, at an average cost of US\$32 per new literate, US\$10 of which came from the United Nations (UNESCO/UNDP, 1976).

vocabularies, ActionAid has gradually dispensed with prepared materials. It uses the techniques of 'participatory rural appraisal' to introduce participants to ways of representing their neighbourhoods and practices symbolically – e.g. maps and calendars – and to progress from there, by way of reflections on their situation, to words and numbers. Literacy is subordinated to empowering the poor to take action to improve their situation; learning sessions are expected to end with identification of action points.<sup>65</sup>

What counts in designing effective programmes? Programmes aimed at helping people acquire the skills they need to sustain a livelihood stand a stronger chance of success than those led by literacy as such;<sup>66</sup> demand is stronger in such cases. Development organizations seem better able than non-specialized educational organizations to design and deliver effective combinations of livelihood and literacy.<sup>67</sup>

Instructors' own skills have a pay-off in terms of skill mastery by learners. The skills most needed in instructors include reliability; mastery of their subject matter, methods and skills; rapport with learners and the ability to sustain learners' interest and engagement.

In many cultures, teaching adults appears to work best when instructor and learners are of the same sex, yet in most countries more men than women are available as instructors while most learners tend to be female. Among the various training models available, recurrent, brief training supports instructors better than initial training followed by irregular supervision. While most people who are literate can be trained to teach literacy effectively to others, livelihood skills require more specialized instructors, preferably selected locally by the learners themselves.

Ideally a group of learners should be served by at least two instructors, each covering selected skills within a coordinated curriculum. Outcomes of adult learning seem to be independent of pedagogical methods: a variety of approaches seems equally effective. That said, it is important for technical materials to both support what the learners wish to learn and to be in advance of what the learners know already. That requires thorough needs assessment, covering both content and qualitative objectives.<sup>68</sup>

Most programmes arrange for learners to undergo some form of assessment and receive certification of their attainments as a means of sustaining motivation. Some recognize the basic literacy course as equivalent to some level of primary schooling, perhaps qualifying learners to enter more formal education.<sup>69</sup> People acquire literacy most easily in their mother language, the next best option being a language with which they are familiar. Ghana has delivered programmes in as many as fifteen languages, and larger countries like India and Nigeria offer even more. Multilingualism, then, is not an insuperable problem even where it adds to programme complexity and cost. Where an international language is the medium for government, education, law and commerce, tuition in that language is sometimes offered as a follow-on to the basic literacy course. Namibia again provides a recent example.

The evidence shows that up to 80% of the people who enrol in well-run literacy classes complete their courses, and half or more of the enrollees pass the local assessment tests. The 'unsuccessful' completers or dropouts may also learn and use some skills. Even better results can be expected from programmes that bring learners together around a common purpose.<sup>70</sup>

The desirable duration of programmes depends on context, but it seems that 400 hours of structured learning can be sufficient to bring totally unschooled and illiterate adults to some basic level of mastery (Oxenham et al., 2002: 38). Similar levels of proficiency have been shown (e.g. in Bangladesh, Indonesia and Uganda) to require some 1,700 to 3,400 hours of primary school instruction (two to four years).<sup>71</sup> This does not imply that adults or schoolchildren are *fully* literate after 400 hours, or two to four years (see UNESCO, 2003a: 60, Figure 2.14) Furthermore, without continued practice, much of what is learned may fall into disuse over the years.<sup>72</sup>

Whether people who attain literacy actually make much use of it is subject to debate.<sup>73</sup> On balance, however, literacy seems more used where economic development is better established.<sup>74</sup> This supports the argument that some degree of economic and political improvement is necessary to sustain literacy (Torres, 2003: 141): people will use their literacy skills where conditions make it useful or desirable for them to do so.

**In many cultures, teaching adults appears to work best when instructor and learners are of the same sex**

65. This strategy is called REFLECT (for 'Regenerated Freirean Literacy through Empowering Community Techniques'). Some 350 organizations in at least sixty countries have adopted it (see [www.reflect-action.org](http://www.reflect-action.org)).

66. See, for instance, Oxenham et al. (2002) and Oxenham (forthcoming).

67. An example is provided by the Farmers' Field Schools in Nepal; see World Education (2000).

68. In a pioneering 1970s example, Indonesia's Paket A project offered 100 booklets on several topics, graded by reading difficulty, from beginners to more fluent readers (Indonesia Ministry of Education and Culture, 1998).

69. In Namibia, literacy programmes are tied in with a national system of vocational qualifications (see Chapter 4).

70. See, for instance, Nirantar (1997) and Oxenham et al. (2002).

71. Bear in mind that literacy is only part of the primary school curriculum.

72. Several studies suggest, however, that the rate of forgetting skills may be faster than had been feared; see, for instance, Kapoor and Roy (1970), Indonesia Ministry of Education and Culture (1998), Cawthera (2003) and Okech et al. (2001).

73. Somewhat negative conclusions are found in Kartekar (2000) on India, and Fiedrich and Jellema (2003) on REFLECT projects in Bangladesh and Uganda. Okech et al. (2001) have more positive observations on other projects in Uganda.

74. See, for instance, Carron, Mwiria and Righa (1989) and Carr-Hill et al. (1991) in their evaluations of literacy programmes in Kenya and Tanzania.

**Better programmes are not necessarily more expensive, from the perspective of society, given the links between learning and income generation**

### Conclusion

Where formal education systems are flanked by programmes of early learning and literacy and skills development, additional benefits accrue to the individual, the community, society, and formal education itself. Children who enjoyed early learning opportunities learn better in formal education, while educated adults, as parents, make bigger efforts to enrol their children and to support them when in school.<sup>75</sup> In this section we have seen that better programmes have a stronger impact on schooling and other developmental objectives than do weak programmes. Furthermore, better programmes are not necessarily more expensive, from the perspective of society, given the links between learning and income generation.

### The quality of schooling

As the two previous sections indicate, qualitative study of country experience and investigation of lessons from particular types of educational programme can bring important insights. This synthetic work, however, is not always as rigorous as is needed to form the basis for educational innovation and policy reform. To establish a better basis for generalization about what policies count, it is necessary to assess the results of a large body of scientific work on what makes a difference for the quality of education. Most of this literature relates to schools, as the major institutions for learning in all societies.

### The paradox

Schooling is a social process, and improvements in resources, technology and the quality of student and teaching inputs should in principle be able to enhance its overall quality. Even a casual look at the history of test scores around the world, however, reveals a central and, at first sight, baffling paradox. In a good number of countries, large increases in average real expenditure per student and other measures of school resources in primary and secondary schools over the last four or five decades have not remotely been matched by a comparable increase in average test scores.

Table 2.4 makes this clear. The first column shows that mathematics and science test scores in eleven OECD countries mainly fell, over the quarter-century ending in 1995,<sup>76</sup> the exception being increases in Sweden, the Netherlands and Italy. Over the same period, however, as the second column shows, six of the eleven countries increased real expenditure per pupil by more than 100%, and the remaining countries also increased spending significantly.

As examples of particular resource changes, PTRs in the United States fell by almost 40% between 1960 and 2000, the proportion of teachers with at least a master's degree doubled and average teacher experience increased similarly. And yet, the mathematics and reading performance of 17-year-old students was only slightly higher in 1999 than it had been 30 years

**Table 2.4: Test scores and changes in real expenditure per pupil, 1970-94 (percentages)**

Country	Change in mathematics and science score, 1970-94	Increase in real spending per pupil, 1970-94	Increase in real GDP per capita, 1970-94	Staff compensation as % of current expenditure on primary education, 1995
Australia	-2.3	269.8	46.4	79
New Zealand	-9.7	222.5	24.3	n.a.
France	-6.6	211.6	60.7	79
Italy	1.3	125.7	74.6	89
Germany	-4.8	108.1	66.8	76
Japan	-1.9	103.3	100.7	87
United Kingdom	-8.2	76.7	58.3	70
Belgium	-4.7	64.7	68	86
Netherlands	1.7	36.3	52.9	78
United States	0	33.1	70.5	80
Sweden	4.3	28.5	35.1	56

75. Oxenham (2004: 8-11), using the Millennium Development Goals as a framework, shows the positive impact of adult learning on each goal, including UPE and gender equality. Myers (2004: 5-6) points to the well-researched benefits of early learning.

76. For the assumptions involved, see Pritchett (2003). Although the calculations are subject to error, this does not affect the broad picture.

Sources: GDP data: constant price chain series (1996 US\$) provided by Heston, Summers and Aten (2002). Other data: Pritchett (2003); OECD (1998: Table B5.1a).

before, when spending was dramatically lower, and the performance of students in science was significantly lower in 1999 than in 1970.<sup>77</sup> As the table shows, the decline was even sharper in some other OECD countries.<sup>78</sup> Despite this evidence, however, popular policy prescriptions for improving school quality focus particularly on further increasing expenditure and resource levels.

## Explanations

Why have countries had such a difficult time increasing average test scores? Some authors have argued that the impact of television on children and their use of time, changes in pedagogy towards less test-oriented teaching methods, the power of teachers' unions and improvements in labour market opportunities for women – allowing highly capable women to move into professions other than teaching – have each had a negative impact on the prospects for learning in schools.<sup>79</sup>

Part of the increase in unit expenditure arises from the growth in societies' wealth since 1970. Column 3 in Table 2.4 shows that all the countries grew richer and, in the last five of those shown, the rise in real GDP per capita was actually greater than the increase in per-pupil spending. This income growth would also have been reflected in individual earnings, and since staff salaries account for 60–90% of unit expenditure (see column 4), the general increase in teacher earnings must account for a significant part of the rise in real expenditure per pupil.

Earnings provide an important incentive mechanism, which can influence both the quality and motivation of teachers. If teachers' real average earnings had kept pace with other professional groups over the period, the productivity impact of their earnings growth would likely have been small. In fact, however, in many countries, teachers' earnings have increased considerably less sharply than those of other groups.<sup>80</sup> So while teachers may actually be better off in real terms than they were in 1970, they may feel worse off, because of their decline in status relative to other professional groups. This circumstance could well explain part – perhaps an important part – of the apparent lack of impact on learning outcomes of increases in real per-student spending over time.

Other things may also make it impossible to observe a strong but simple relationship between resource inputs to schooling and test scores. Changes in other inputs (teaching time, curricula requirements, school sizes, etc.) could have affected test scores as well. The second half of the twentieth century was a period of strong expansion of school systems throughout the world. In many countries the original social roles of schooling have changed, yet expectations have been slower to adjust as a greater mix of ability groups has proceeded through to the higher levels of schooling.

Many studies have been conducted in the past twenty years to discover the extent to which such disturbances explain why the expected relationship between inputs and outputs is concealed. In the economics literature there have been two main types of investigation. The first involves cross-country studies of the relationship between educational expenditure and test scores. The second, now comprising numerous articles and papers, takes a more micro approach and is based on the idea of an education production function to inform the analysis. The results of both types of work are briefly discussed below.

## Results from macro studies

Table 2.5 summarizes the main results of the small number of empirical papers that has sought to establish whether there is a positive, strong and significant causal relationship between educational expenditure and outcomes at the aggregate level. The table shows the dependent variables that were taken as proxies for school quantity or quality in each study: test scores, repetition and dropout rates, completion rates, enrolment ratios at primary, and sometimes secondary, level. The studies aim to establish the extent to which increases in school resources – usually measured as PTRs, expenditure per pupil, proportion of GDP or average teacher salaries – enhance educational outcomes.

The strongest set of results seeming to establish the importance of school resources for test scores emerges from the work of Lee and Barro (2001). They find, first, that increases in PTR result in a decrease in average test scores, thereby indicating that smaller classes are associated with better pupil achievement; their

**The second half of the twentieth century was a period of strong expansion of school systems throughout the world**

77. Expenditure and input figures are from US Department of Education data, while scores are the results of tests administered to a random sample of students aged 17 by the National Assessment of Educational Progress. See Hanushek, 2003: 67–9).

78. See also Gundlach, Woessmann and Gmelin (2001) and Carnoy (2004).

79. Pritchett (2003). See also Carnoy (2004).

80. See the section on teacher earnings in Chapter 4.

**Table 2.5: Results of studies investigating the relationship between educational expenditures and outcomes**

Study	Type of data, year, number of observations	Educational outcome	Educational expenditure	Results
Al-Samarrai (2002)	Cross-country; 1996; 33 to 90	<ul style="list-style-type: none"> <li>Primary gross and net enrolment ratios</li> <li>Survival rate to primary grade 5</li> <li>Primary school completion rate</li> </ul>	<ul style="list-style-type: none"> <li>Public primary education spending (% GNP)</li> <li>Primary expenditure per pupil</li> <li>Primary PTR</li> </ul>	<ul style="list-style-type: none"> <li>Primary expenditure per pupil has a positive and significant impact on survival rate to grade five only (10%), and a negative and significant impact on the enrolment ratios (1% and 5%, respectively).</li> <li>All other coefficients are insignificant.</li> </ul>
Hanushek and Kimko (2000)	Cross-country panel; 1965, 1970, 1988, 1991; 67 to 70	<ul style="list-style-type: none"> <li>IEA and IAEP<sup>1</sup> mathematics and science tests</li> </ul>	<ul style="list-style-type: none"> <li>PTR</li> <li>Current education spending per pupil</li> <li>Total expenditure on education (% GDP)</li> </ul>	<ul style="list-style-type: none"> <li>Insignificant</li> <li>Negative impact, 1%</li> <li>Negative impact, 5%</li> </ul>
Woessmann (2000)	Cross-country; 1995; 39	<ul style="list-style-type: none"> <li>TIMSS mathematics and science scores</li> </ul>	<ul style="list-style-type: none"> <li>Class size</li> </ul>	<ul style="list-style-type: none"> <li>Positive impact, 1%</li> </ul>
Lee and Barro (2001)	Cross-country panel: 1964, 1970, 1975, 1980, 1982, 1984, 1985, 1990 (depending on the outcome considered) 214 to 346	<ul style="list-style-type: none"> <li>Test scores</li> <li>Primary school repetition rates</li> <li>Primary school dropout rates</li> </ul>	<ul style="list-style-type: none"> <li>PTR</li> <li>Average teacher salary</li> <li>Current education spending per pupil</li> </ul>	<ul style="list-style-type: none"> <li>PTR has a negative and significant impact on all three variables (5%, 1% and 5%).</li> <li>Average teacher salary has a positive and significant impact on test scores (10%).</li> <li>All other coefficients insignificant.</li> </ul>
McMahon (1999)	Cross-country; early 1990s; 44 to 50	<ul style="list-style-type: none"> <li>Primary male and female gross enrolment ratios</li> <li>Male and female grade 5 completion rate</li> </ul>	<ul style="list-style-type: none"> <li>Public current expenditure on primary education (% GNP)</li> <li>Public current expenditure per primary student (% GNP per capita)</li> <li>Public current expenditure per primary student (level)</li> </ul>	<ul style="list-style-type: none"> <li>Total expenditure has a positive and significant impact (1%).</li> <li>Per capita expenditure has a negative and significant impact (1%).</li> <li>Positive impact, 1%</li> </ul>
Gupta, Verhoeven and Tiongson (1999)	Cross-country: 1993/94 23 to 42	<ul style="list-style-type: none"> <li>Primary and secondary gross enrolment ratios</li> <li>Persistence through grade 4</li> </ul>	<ul style="list-style-type: none"> <li>Primary and secondary education spending (% total education spending)</li> <li>Education spending (% GDP)</li> </ul>	<ul style="list-style-type: none"> <li>Expenditure on primary and secondary education has a positive and significant impact on all three variables (1–5%, 5–10%, 5–10%)</li> <li>Total education spending has a significant and positive impact on secondary enrolment only (5%); insignificant for the other two variables.</li> </ul>
Schultz (1996)	Country panel; 1965–80; 60 to 191	<ul style="list-style-type: none"> <li>Primary gross enrolment ratio</li> </ul>	<ul style="list-style-type: none"> <li>Public teacher compensation as (% GNP per working age adult)</li> </ul>	<ul style="list-style-type: none"> <li>Negative impact, 1%</li> </ul>
Colclough with Lewin (1993)	Cross-country; 1986; 82	<ul style="list-style-type: none"> <li>Primary gross enrolment ratio</li> </ul>	<ul style="list-style-type: none"> <li>Public current expenditure on primary (% GNP)</li> <li>Public current expenditure per primary student (% GNP per capita)</li> </ul>	<ul style="list-style-type: none"> <li>Insignificant</li> <li>Negative impact, 1–5%</li> </ul>

*Notes and sources:* Table: Al-Samarrai (2002), with additional results from Gupta, Verhoeven and Tiongson (1999). Detailed sources: Hanushek and Kimko (2000: Table 3); Woessmann (2000: Table 1); Lee and Barro (2001: Table 3); McMahon (1999: 164 and 166); Schultz (1996: Tables 2 and 3); Colclough with Lewin (1993: Table 2.6a); Gupta, Verhoeven and Tiongson (1999: Table 1); Al-Samarrai (2002: Tables 4.1 and 4.2). Lee and Barro (2001) present other specifications but the results do not differ markedly.

Colclough with Lewin (1993) also present results for developing countries and African countries separately, but the results on the resource variables are similar.

1. International Assessment of Educational Progress.

results suggest that a decrease in the average PTR (by 12.3 in 1990) raises test scores by 1.8 percentage points. Second, higher teacher salaries in their sample of fifty-eight countries were associated with a significant increase in test scores. Third, the results for repetition and dropout rates are consistent with their being affected by school resources: reductions in PTR are associated with reduced rates of dropout and

repetition. Lee and Barro conclude that school inputs (especially smaller class sizes, but probably also higher teacher salaries and more years of schooling) enhance educational outcomes.

Two studies produce evidence that spending on the primary system – but not total education spending – affects the retention rate to grades 4

and 5 (Gupta, Verhoeven and Tiongson, 1999; McMahon, 1999). Not surprisingly, it is the share of resources devoted to primary schools that affects their functioning, total resources going to education being too approximate a measure to be relevant.

Other results, however, are rather different. As the table shows, Hanushek and Kimko (2000) found that variations in school resources do not have strong effects on test performance. By their estimates, the effects of various types of resources are either statistically insignificant or, more frequently, statistically significant but with an unexpected sign – in other words, suggesting that increased resources actually results in lower test scores. This finding holds regardless of the measure of resources used: PTR, current expenditure per student, total expenditure per student and a variety of others not shown in the table.

A later study (Hanushek and Luque, 2003) uses data from the TIMSS, taken from representative samples of schools in 1995 in more than forty countries for pupils aged 9, 13 and 17. The paper examines the determinants of pupil achievement, focusing on the class-level averages of mathematics test results for 9- and 13-year-olds. It finds that almost all the coefficients associated with school characteristics are insignificant and divided almost equally between negative and positive ones. Only five countries exhibit the 'expected' negative and significant relationship between achievement and class size; none exhibits a positive and significant relationship between achievement and teacher education or experience. The authors conclude that, overall, for the countries sampled, resources' power to produce better student performance is rather limited. While certain countries do stand out for having significant effects, these results do not suggest that outcomes related to school resource differences are more positive in poorer countries or in those that begin with lower levels of resources.

The other studies reported in the table deal with educational outcomes expressed in terms of enrolment ratios, rather than qualitative indicators. Here, too, the results conflict. Some studies consistently find that expenditure increases are associated with increased enrolments, and others find the opposite.

In general, no consistent relationship is discernable across these studies. Those using internationally comparable test scores tend to show that resources have a significant impact on educational outcomes, but the direction of the impact differs among the studies and is often counter-intuitive. Taken as a whole, this group of studies does not suggest that a positive, strong and significant relationship between educational expenditure and outcomes can yet be identified using aggregate, country-level data.

### Results from micro studies

Addressing questions about the relative importance of schools and families in determining cognitive achievement and subsequent success in life is not a new activity. For example, a century ago French sociologist Paul Lapie studied the life histories of 722 men who had been enrolled between 1872 and 1893 at the primary school in Ay, a small town in eastern France. He concluded, on the basis of this work: 'Schooling sometimes succeeds in breaking the strings of the net by which economic circumstances control our destinies. Its impact is not great, but it is not nothing.'<sup>81</sup> The notion that education in general, and school in particular, makes a difference to people's lives informed politics and public policy throughout the 20th century. Research on the matter was given a strong stimulus in the mid-1960s by the publication of an influential report concluding that, in the USA, family background and the composition of peer groups in school had a much larger impact on educational outcomes and on subsequent economic success than did variations in the characteristics of the schools themselves (Coleman et al., 1966). The profoundly radical implication was that schools simply helped reproduce inequalities in society that already existed. One result was an outpouring of research aimed at establishing a firmer basis for these findings or refuting them.

This modern economic approach to investigating the determinants of educational outcomes has borrowed well-established techniques from other economic applications. The idea that there is a determinate relationship between inputs to a production process and the outputs that subsequently emerge has long been important in microeconomic analysis. If a firm's production possibilities are governed by certain technical

**The profoundly radical implication of the Coleman report was that schools simply helped reproduce inequalities in society that already existed**

81. 'Ainsi, l'école réussit parfois à rompre les mailles du réseau dans lequel des causes d'ordre économique enferment nos destinées. Son action n'est pas considérable, mais elle n'est pas nulle'. Paul Lapie. 1904. 'Les Effets sociaux de l'école' ['The Social Effects of Schooling'], *La Revue scientifique* (Revue rose), 41 (2); cited in Baudelot et al. (2004).

**The application of production function analysis to education is somewhat hazardous**

relationships between factors of production – e.g. it takes a bag of nails, some tools, some planks of wood and some days of labour to produce a fence – the ‘production function’ describes the maximum feasible output (fence) obtainable from alternative combinations of these inputs. Production functions are powerful analytic tools, which have been applied to the analysis of most forms of economic production. Since the mid-1960s, they have also been widely used in the economic analysis of education.

The application of this idea to education is, however, somewhat hazardous. While it may be reasonable to assume that managers of firms in some sense ‘know’ the shape and characteristics of the production function they face in their industry, this is not the case with education managers. Furthermore, the inputs to the schooling process are much less homogeneous (teachers, goals, pupils) than in industry, where labour usually faces more defined tasks than do teachers. Nor can the characteristics of the outputs (more schooled pupils) be unambiguously compared with earlier pupil inputs for ‘value added’ purposes. But the main difficulty with representing education as a production process is that some of its inputs and all of its outcomes are embodied in pupils, who have their own autonomous behaviour. Planks of wood cannot decide that they do not want to be assembled, avoid coming to the construction site or refuse to interact with construction workers. Unfortunately, the production function literature has comprehensively failed to model pupil behaviour.

Notwithstanding these and other difficulties,<sup>82</sup> this literature has burgeoned, with the explicit aim of guiding education policies around the world by identifying which policy-controlled inputs have the largest marginal impact on achievement. Would a greater impact on learning, for example, arise from providing more textbooks, increasing teacher salaries or improving teacher training? In the absence of proper theories of cognitive processes, pupil-teacher interactions and schools as institutions, the usual assumption has been that higher expenditure and input provision will automatically lead to better educational outcomes. From what has already been said at the macro level, it is perhaps not surprising that such a clear picture does not emerge from the research results at the micro level.

**Results from education production function studies**

Since the education production function literature now includes numerous studies, a summary analysis of this evidence might be expected to provide robust empirical regularities as to which inputs stimulate achievement, and to what extent.<sup>83</sup> Table 2.6 shows a meta-analysis of all published estimates up to 1995 for the USA. It covers eighty-nine individual publications containing 376 production function estimates. The educational outcomes considered in three-quarters of these studies are test scores; the rest used other outcomes, such as rates of continuation in school or drop-out rates. Resources considered are real classroom resources (PTR, teacher education, teacher experience), financial aggregates (teacher salary, expenditure per pupil) and other resources (facilities, administration, teacher test scores).

The results show that for between 9% and 37% of these cases, resource inputs had a positive effect upon measured student achievement. Their frequency was often not much greater than the number of negative cases, however: most notably, improving PTR had a negative impact upon achievement as often as it had a positive one.<sup>84</sup> Furthermore, although the proportion of positive and significant estimates was, on the whole, much greater than that of the negative ones, in a considerable majority of all cases there was no significant measured relationship between resource inputs and student performance.

The developing-country sample, shown in Table 2.7, summarizes ninety-six estimates of the impact on student performance of up to six resource inputs (PTR; teacher education, experience and salary; expenditure per pupil; and school facilities). Here, there are twelve to sixty-three estimates per variable, and the results are much more varied than those derived from the US studies. First, there are proportionately fewer statistically insignificant estimates: in only two cases do these represent a majority. Second, among the significant estimates, a strong majority are positive, and the proportion of negative estimates is below 10% except in the cases of the PTR and teacher salary. Third, an absolute majority of the estimates are significant and positive for teacher education, expenditure per pupil and school facilities.

82. See Glewwe (2002) for discussion of the technical problems underlying the methodology.

83. It does so if one accepts the implicit assumption that a homogenous ‘technology’ exists, whose parameters would be common to all education systems across the world, conditional on a sufficient number of adequately selected control variables.

84. The regression coefficient may be of the opposite sign, e.g. a negative coefficient associated with the PTR is considered a positive estimate, for it implies that teacher resources are positively associated with achievement.

The literature contains much debate as to the implications of all these research results. Some argue that the glass is half full, others that it is half empty. Alternative ways of grouping and summarizing the results, however, suggest that improvements to the resources available to developing country schools are more likely than not to improve the levels of cognitive achievement of pupils.<sup>85</sup> This suggests grounds for confidence that, in developing countries, increased resources for schools do influence student performance positively – more strongly than in the case of the USA, with its already high average resource levels.

That being said, dozens of production function studies for education have been published over the last three decades for the USA, a few other developed countries and a relatively large number of developing countries. They have yielded few uncontroversial empirical regularities on which education policies can be based. Quite apart from the considerable technical problems involved in interpreting individual estimates, there is no real consensus on whether it makes sense to summarize them through meta-analytical techniques (as in Tables 2.6 and 2.7) rather than surveying a few key studies, nor on how the meta-analysis should be conducted and its results interpreted. Given this level of uncertainty, at least some of the conclusions that the various authors in this field have drawn depend on their *a priori* opinions, as much as anything, and cannot be unquestioningly used to guide policies.

### Effective schools research

In parallel with this economic tradition, a rather different empirical approach to the study of schools and classrooms began to emerge. Education researchers became increasingly concerned that standard production function approaches ignored important aspects of the processes of learning and teaching in schools. These approaches tended to treat what happened in schools – the quality and nature of teacher-pupil interactions, the ways resource inputs were actually used – as being of little consequence. Yet it appeared obvious that ways of organizing schooling – e.g. time spent in class, amount of homework, ways of assessing pupil progress, and teacher expectations, experience and in-service training – were likely to affect student

**Table 2.6: Percentage distribution of estimated effect of key resources on student performance, based on 376 production function estimates (United States)**

Resources	Number of estimates	Statistically significant (%)		Statistically insignificant (%)
		Positive	Negative	
<b>Real classroom resources</b>				
PTR	276	14	14	72
Teacher education	170	9	5	86
Teacher experience	206	29	5	66
<b>Financial aggregates</b>				
Teacher salary	118	20	7	73
Expenditure per pupil	163	27	7	66
<b>Other</b>				
Facilities	91	9	5	86
Administration	75	12	5	83
Teacher test scores	41	37	10	53

Source: Hanushek (2003a), based on Hanushek (1997).

**Table 2.7: Percentage distribution of estimated expenditure parameter coefficients from ninety-six education production function estimates (developing countries)**

Inputs	Number of estimates	Statistically significant (%)		Statistically insignificant (%)
		Positive	Negative	
PTR	30	27	27	46
Teacher education	63	56	3	41
Teacher experience	46	35	4	61
Teacher salary	13	31	15	54
Expenditure per pupil	12	50	0	50
Facilities	34	65	9	26

Source: Hanushek (2003a), based on Hanushek (1995).

outcomes quite apart from the resource inputs to schooling, and that this might account for the often ambiguous results of production function studies.

The studies in this emerging 'school effectiveness' approach<sup>86</sup> remained quantitative in orientation, and mainly focused on the school as the unit of analysis. Generally, too, the concern was to explain pupils' academic outcomes. Their main innovation was to add new school-level factors to the earlier analyses that had looked at processes within schools only to the extent that they were believed to be important for academic outcomes. Taking a broader perspective, it became increasingly clear that different patterns of school organization and teacher behaviour had different effects on

85. The approach used in Tables 2.6 and 2.7 can be challenged on the grounds that giving equal weight to all estimates made in each study, as opposed to counting only the major results, of each, biases the results, because studies with many estimates are likely to include a greater proportion of negative or insignificant relationships. Krueger (2003) finds that the incidence of studies with positive results is twice as high as the incidence of those with negative results, and that the probability of this occurring by chance is very small. Accordingly, marginal increases in resource inputs, on this interpretation of the production function evidence, are very important for improving cognitive achievement.

86. Brookover et al. (1979) and Mortimore et al. (1988) were early examples.

learning outcomes.<sup>87</sup> Results for industrialized countries converged around the importance of five factors:

- strong educational leadership;
- emphasis on acquiring basic skills;
- an orderly and secure environment;
- high expectations of pupil attainment;
- frequent assessment of pupil progress.

Most of the many reviews of this large body of literature focus on 'what works'.<sup>88</sup> Table 2.8 summarizes results from five such reviews.

Although the emphasis in these studies varies, consensus is strong with respect to the aforementioned five factors (represented in the first five rows of the table). Other factors

increasingly being found to be important are school-based in-service development of teachers, the use of well-structured, purposeful and sustained teaching time and the beneficial influence of parental support and involvement with the school.

On the other hand, the amount of variation in pupil achievement that these empirical studies in industrialized countries typically explain is comparatively low, seldom more than 15%. Although this phenomenon is partly due to measurement error, to the lack of an underlying theoretical model that can effectively capture all the influences upon student learning and to the lack of direct value-added measures in most studies, the power of school and classroom variables to explain achievement remains limited.

**Table 2.8: Effectiveness-enhancing conditions of schooling: results of five review studies**

Purkey & Smith, 1983	Levine & Lezotte, 1990	Scheerens, 1992	Cotton, 1995	Sammons, Hillman & Mortimore, 1995
■ Strong leadership	■ Outstanding leadership	■ Educational leadership	■ School management and organization, leadership and school improvement, leadership and planning	■ Professional leadership
■ Clear goals on basic skills	■ Focus on central learning skills		■ Planning and learning goals and school-wide emphasis on learning	■ Concentration on teaching and learning
■ Orderly climate; achievement-oriented policy; cooperative atmosphere	■ Productive climate and culture	■ Pressure to achieve, consensus, cooperative planning, orderly atmosphere	■ Planning and learning goals, curriculum planning and development	■ Shared vision and goals, a learning environment, positive reinforcement
■ High expectations	■ High expectations		■ Strong teacher-student interaction	■ High expectations
■ Frequent evaluation	■ Appropriate monitoring	■ Evaluative potential of the school, monitoring of pupils' progress	■ Assessment (district, school, classroom level)	■ Monitoring progress
■ Time on task, reinforcement, streaming	■ Effective instructional arrangements	■ Structured teaching, effective learning time, opportunity to learn	■ Classroom management and organization, instruction	■ Purposeful teaching
■ In-service training/ staff development	■ Practice-oriented staff development		■ Professional development and collegial learning	■ A learning organization
	■ Salient parental involvement	■ Parent support	■ Parent-community involvement	■ Home-school partnership
		■ External stimuli to make schools effective ■ Physical and material school characteristics ■ Teacher experience ■ School context characteristics	■ Distinct school interactions ■ Equity ■ Special programmes	■ Pupil rights and responsibilities

87. A related body of literature on 'school improvement' is mainly concerned with praxis, adopting a particular set of tenets about how to improve conditions in schools. In that sense it represents more a set of policy choices than an established body of knowledge about how schools improve. Chapter 4 discusses some implications of this literature.

88. Numerous reviews on school effectiveness have been published since the late 1970s. Early reviews are those by Anderson (1982), Cohen (1982), Dougherty (1981), Edmonds (1979), Murnane (1981), Neufeld, Farrar and Miles (1983), Purkey and Smith (1983), Rutter (1983), Good and Brophy (1986), Ralph and Fennessey (1983), Kyle (1985) and Sweeney (1982). More recent reviews are those by Levine and Lezotte (1990), Scheerens (1992), Creemers (1994), Reynolds, Hopkins and Stoll (1993), Sammons, Hillman and Mortimore (1995) and Cotton (1995). A comprehensive overview of methodology and results is provided in Teddlie and Reynolds (2000).

Of course, classroom resources typically vary less in rich countries than in developing countries. The impact of class size on pupil learning when the range is fifteen to thirty-five students is different than when the upper limit is as high as 100, as may occur in many developing countries. Studies do seem to suggest that school-related factors explain more of the variation in achievement in developing countries than in industrialized countries, that the impact of socio-economic background is less in the former and that the factors with the greatest importance for student achievement appear to be somewhat different. A review of research on the factors promoting science achievement in developing countries found that in over four-fifths of the studies the length of instructional programmes, use of a school library and school meals were important factors, and that over two-thirds identified teacher training and the presence of textbooks and instructional materials as important (Walberg, 1991). By contrast, only one-quarter to one-third of the studies found the presence of science laboratories, increases in teacher salaries and reductions in class size to be important.

This evidence suggests that, as we would expect, resources are more important determinants of pupil achievement in resource-poor environments than in richer ones. The importance of teacher quality also tends to emerge more clearly, no doubt because the variation in levels and quality of teacher training, competence and initial educational background is much generally greater in developing countries.

#### **Evidence on instructional effectiveness**

School resources provide a framework within which teachers can guide students in their learning. In a more fundamental sense, however, the effectiveness of learning depends on what pupils, as well as teachers, bring to the task. Early attempts to understand the learning process emphasized both aspects. It became clear that the following five elements were likely to have an important influence on learning outcomes (Carroll, 1963 and 1989):

- *aptitude*, which determines the amount of time a pupil or student needs to learn a given task under optimal conditions of instruction and student motivation;

- *opportunity to learn*, measured as the amount of time available for learning;

- *perseverance*, or the amount of time a pupil or student is willing to spend on learning;

- *quality of instruction*, which, when sub-optimal, increases the time needed for learning;

- *ability to understand instruction*, which includes language comprehension and the ability to understand the nature of the task and how to go about it.

Although the above formulation is not specific as to what determines high-quality instruction, teachers have long been believed to be the key to education quality. Accordingly, education researchers have focused extensively on the best ways to improve teacher effectiveness. Early work found little consistency in the relationships between teachers' personal characteristics (e.g. warm and open or strict and inflexible) and pupil achievement (Medley and Mitzel, 1963; Rosenshine and Furst, 1973; Gage, 1965). As later research focused on classroom observation of teacher behaviour, more systematic links between some teacher characteristics (e.g. clarity, flexibility, enthusiasm, ordered preparation) and pupil performance were suggested. The research methodologies remained controversial (Weeda, 1986), however, and no decisive commonalities had yet emerged.

Over the 1980s, studies of the determinants of effective instruction began to show more consistent results. It seemed that, at primary level, effective learning time, class organization and management, teaching strategies and instruction, and assessments and teachers' expectations were all significant factors in improving pupil performance (Stallings, 1985). As regards *effective learning time*, it became clear that simply making the school day longer did not necessarily lead to better performance. More important, ultimately, is how effectively time is spent. Studies of effective teachers (Stallings and Mohlman, 1981) showed that they spent some 15% of the school day on organization, management and lesson planning; 50% on interactive teaching and 35% on monitoring pupils' work.

**Resources are more important determinants of pupil achievement in resource-poor environments than in richer ones**

**High teacher expectations contribute significantly to pupil performance**

As regards *class organisation and management*, streaming on the basis of ability appears to work for more gifted pupils, but evidence from a large number of surveys indicated hardly any effect for less able groups (Kulik and Kulik, 1982; van Laarhoven and de Vries, 1987; Reezigt, 1993; Slavin, 1987). Not surprisingly, in classes where there is disruptive behaviour, effective learning time is reduced and pupil performance is lower.

Appropriate *teaching strategies* depend upon the type of learning tasks targeted. For example, those largely depending on memory are most effectively taught with a highly ordered and consistent approach. For acquisition of new understanding, a clear presentation of the information is crucial, as are questions to check whether pupils have absorbed specific insights. With regard to problem solving, evidence suggests that it is desirable for pupils to take much of the initiative.<sup>89</sup>

Studies on *teacher assessments and teachers' expectations* of pupils underline the danger of self-fulfilling prophecies. Where teachers form negative expectations of certain pupils, they are likely to give them less attention and expose them less to more challenging tasks. This kind of stereotyping can have highly negative consequences for some pupils. (Hoeven-Van Doornum and Jungbluth, 1987). High teacher expectations, on the other hand, contribute significantly to pupil performance.

A typical list of steps/characteristics required for effective teaching, based on these and other research results, is the following (in this case, from Doyle, 1985):

- clearly formulated teaching goals;
- course material carefully split into sequenced learning tasks;
- clear explanations from the teacher as to what pupils must learn;
- regular questioning by teachers to gauge pupils' progress and whether they have understood material covered;
- ample time for pupils to practice what has been taught, with much use of 'prompts' and feedback;

- skills being taught until mastery of them is automatic;

- regular testing, and expectations that pupils will be accountable for their work.

This highly structured approach to teaching<sup>90</sup> appears to work equally well for primary and secondary education. Of course, at secondary level, the range of subject options is greater, children are older and a stronger emphasis is needed on more advanced cognitive processes (Brophy and Good, 1986: 367). Furthermore, progress through the subject matter can be quicker, and testing need not be so frequent. The importance of varying learning tasks and of providing intellectual challenges is clear. An evaluative climate in the classroom, with pupils encouraged to take risks even in complicated tasks, is helpful. The evidence on the impact of individual teaching and of working together in small groups is mixed. In general, it is difficult to point to clear evidence of higher achievement deriving from either mode.<sup>91</sup> It seems, however, that group work can be beneficial where group rewards are introduced, based on the individual learning of all group members, and where students are taught how to work together in ways that closely reflect their instructional objectives (Slavin, 1996: 57).

Not all education analysts, however, accept the value of the notion of structured teaching. Constructivists, for example, believe learners should be the main instigators and designers of learning processes. Box 2.3, Table 2.9 contrasts some major distinguishing features of learning and instruction according to the constructivist position with characteristics of more traditional models. Such comparisons, of course, risk oversimplification. Furthermore, less extreme constructivist views can be reconciled with other approaches (Merrill, 1991), for example, when teacher-controlled and learner-controlled instructional situations are used alternately (Boekaerts and Simons, 1993). However, it is difficult to avoid the conclusion that such approaches to teaching presuppose levels of student competence and classroom facilities that may not be compatible with low-resource environments, where many pupils come from poor or non-literate backgrounds. This debate may therefore be particularly relevant to the circumstances of richer nations and communities.

89. Collins and Stevens (1982) list five teaching strategies to support problem solving: a) systematic variation of examples; b) counterexamples; c) entrapment strategies; d) hypothesis identification strategies; and e) hypothesis evaluation strategies.

90. See also Bloom (1968) on the notion of 'mastery learning' and Rosenshine (1983) on 'direct instruction'; both suggest characteristics of effective teaching that are quite similar to this list.

91. Bangert, Kulik and Kulik (1983) found individualized teaching in secondary education had little or no impact on achievement or on pupil self-esteem and attitudes. Vedder (1985) suggested that the lack of a clear positive influence of group work was due to the way pupils work together, there being insufficient cognitive stimulation in such circumstances.

**Box 2.3 Constructivism**

An important group of analysts stress the key role of the learner rather more than that of the teacher. These ‘constructivists’ argue that learning needs to be largely self-directed, with an emphasis on discovery, and with particular value being ascribed to the students’ interpretation of events and facts. Learning strategies, and reflection upon them, are as important as mastering content. Recognizing different ways of finding a solution is as important as the solution itself. Terms like ‘active learning’ (Cohen, 1988), ‘situated cognition’ (Resnick, 1987) and ‘cognitive apprenticeship’ (Collins, Brown and Newman, 1989) are used.

Consistent with this approach, teaching strategies need to enable students ‘to construct their own meaningful and conceptually functional representations of the external world’ (Duffy and Jonassen, 1992: 11). The teacher becomes more of a coach, who assists students in ‘criss-crossing the landscape of contexts’ (Spiro et al., 1992: 8). Some have used the term ‘adventurous teaching’ to describe this approach (Cohen, 1988).

There is less emphasis here on prior structuring of goals, learning tasks and plans; goals are supposed to emerge when ‘situated learning’

takes place and plans are not so much supposed to be submitted to the learner as constructed in response to changing demands and opportunities. Learning situations should encourage students to engage in sustained exploration. Some authors writing from this perspective state that ‘transfer’ is the most distinguishing feature (Tobias, 1991), whereas others mention argument, discussion and debate to arrive at ‘socially constructed meaning’ (see Cunningham, 1991).

In this context, the role of assessment and the evaluation of students’ progress is hotly debated. Radical constructivists take the position that performance on an actual learning task is the only legitimate form of assessment, since distinct ‘external’ evaluation procedures could not do justice to the specific meaning of a particular learning experience for the student. Others (e.g. Jonassen, 1992) judge that assessment procedures should merely be different: goal-free, rather than fixed on particular objectives, formative rather than summative and oriented towards assessing learning processes rather than mastery of subject matter. Appraisals of samples of products and portfolios qualify as acceptable assessment procedures.

**Classroom practices have the strongest association with achievement**

Research evidence has continued to accumulate. A review of the influence of educational, psychological and social factors on learning (Wang, Haertel and Walberg, 1993) ranks factors that have been found to influence student achievement as follows (from high to low):

- student characteristics;
- classroom practices;
- home and community educational contexts;
- design and delivery of curriculum and instruction;
- school demographics, culture, climate, policies and practices;
- state and district governance and organization.

Apart from the characteristics of the students themselves, classroom practices have the strongest association with achievement. Within this category are classroom management and student and teacher social interaction (e.g. ‘students respond positively to questions from

**Table 2.9: Comparison of traditional and constructivist instructional models**

Traditional instruction	Instruction inspired by constructivism
Emphasis on basic skills	Bias towards higher-order skills
Subject matter orientation	Emphasis on learning process
Structured approach <ul style="list-style-type: none"> <li>■ pre-specified objectives</li> <li>■ small steps</li> <li>■ frequent questioning/feedback</li> <li>■ reinforcement through high level of mastery</li> </ul>	Discovery-learning <ul style="list-style-type: none"> <li>■ ‘rich’ learning environment</li> <li>■ intrinsic motivation</li> <li>■ challenging problems</li> </ul>
Abstract-generalizable knowledge	Situation-specific knowledge, learning from cases
Standardized achievement tests	Assessment; less circumscribed alternative procedures

Source: Scheerens (2004)

other students and from the teacher’). An illustrative variable within classroom management is ‘group alerting’ (teachers using

**Grouping for reading across grades appears to have a consistently positive effect**

questioning/recitation strategies that maintain active participation by all students). Other relatively influential aspects of classroom practice are classroom climate, classroom assessment and quantity of instruction (e.g. time on task). In summary, it appears that, as far as classroom instruction is concerned, general instructional approaches that are relatively structured bring important pay-offs for learning.

One study (Slavin, 1996), using a combination of systematic literature review and meta-analysis, has analyzed research in which pupil outcomes among those grouped according to ability were compared with those from heterogeneously grouped classes. It concludes that *whole-class ability grouping*, with the aim of creating groups that are homogeneous in ability level ('ability-grouped class assignment'), is generally ineffective: in the studies surveyed, effects were either negative or close to zero, with only a few cases reporting small positive effects. On the other hand, grouping for reading across grades appears to have a consistently positive effect, and it appears that *within-class ability grouping in mathematics* at primary level is also effective.<sup>92</sup> These results suggest that successful ability grouping requires:

- reduced heterogeneity of aptitudes for the skill being taught;
- mechanisms to deal with errors in assignment to ability groups and changes in student performance;
- teachers varying their pace and level of instruction to correspond to students' levels of readiness and ability to learn.

Ability-grouped class assignment, regardless of the skills, probably fails the first and often fails the second of these criteria (Slavin, 1996: 158).

Within-class ability grouping entails some loss of teaching time during transition of the teacher from one group to the next. It also implies that groups will have to work a considerable amount of time without direct teacher instruction. Nevertheless, the above results indicate that within-class ability grouping is more effective despite a certain loss of direct instructional time. It would seem advisable, then, not to allow too many ability groups within classes, to minimize

the negative impact of such loss. Other implications are (Slavin, 1996: 164):

- students should be assigned to heterogeneous classes most of the time and be regrouped by ability only in subjects in which low heterogeneity is particularly important (e.g. reading, mathematics);
- grouping plans should reduce heterogeneity in the skill being taught;
- grouping plans should be flexible and allow for easy reassignment;
- teachers should vary their level and pace of instruction so as to be consistent with students' levels of performance;
- the number of groups should be minimized.

More recent contributions to the study of instructional effectiveness have returned to examination of the personal characteristics of effective teachers. They reaffirm the importance of subject matter mastery and verbal skills (Darling-Hammond, 2000) and the strong impact on pupil performance of teachers' expectations, their drive for improvement amongst their pupils and their passion for learning (McBer, 2000). Another emerging theme is the importance of structured teaching mixed with respect for the self-regulated learning of pupils (Brophy, 2001; Baumert, Blum and Neubrand, 2000; Anderson, 2004).

**Implications for the major determinants of effective teaching**

If learning outcomes depend on student learning strategies and students' motivation to learn (engagement), the teaching conditions summarized in Table 2.10 are also important, particularly the top three: relevance, time and structure.

First, relevance includes keeping subject matter selection well aligned to the intended curriculum – including between grades and classes – and ensuring that the contents of teaching and learning assignments match those of tests and other assessment instruments.

The second condition requires attention to the time to be spent on major curriculum areas and

92. The median sizes for these effects were + 0.44 standard deviation for reading and +0.34 standard deviation for mathematics.

**Table 2.10: The most important conditions for enhancing teaching effectiveness**

Relevance	Curriculum alignment
Time	Learning time
Structure	Structured teaching Stimulating engagement Monitoring and questioning Feedback and reinforcement Modelling learning/self-regulation
Classroom environment	Task-oriented climate Mutual respect Orderliness, safety
Teacher characteristics	Subject matter mastery Verbal intelligence Teaching repertoire Achievement orientation

Source: Scheerens (2004)

subjects according to official timetables. At school level, 'net teaching time' in the classroom – i.e. official teaching time minus time 'lost' to other activities – is a key quality variable. Ultimately, optimizing the time students spend actively involved in learning activities has a significant impact on pupil achievement.

Third, structured teaching emerges as being important to pupil performance in many instructional effectiveness studies, perhaps most strongly for less-able learners in primary schools, but also more generally at higher levels of schooling where more advanced cognitive skills are targeted. Debates inspired by constructivism suggest a continuum running from a highly structured teaching process to one characterized by a high level of learner independence. The key to effective teaching lies in selecting the appropriate extent of structure, in the light of learner characteristics, learning tasks and educational objectives. Structural dimensions also include frequent monitoring of students' progress, provision of feedback and reinforcement related to assessment outcomes, provision of cognitive support and fostering of student engagement.

Fourth, classroom climate and organization, the extent of mutual trust, and safety and discipline factors can have a direct impact on student learning. Fifth, strong teacher characteristics are prerequisites for delivering relevant teaching, using time efficiently and

providing an appropriate level of structure. As discussed earlier, broader conditions for effective schooling, such as school organization and educational leadership, are necessary supportive conditions for effective teaching.

### The experimental evaluation of education policy interventions

The difficulty of estimating education production functions has led to the development of a new approach to measurement of the impact of school resources on education outcomes: the experimental evaluation of policy interventions. Rather than seeking to identify the parameters of a theoretical model of education production that might have universal relevance, this approach attempts to measure the impact of a given change in school resources on education outcomes in a specific institutional context. If the former model was empirical testing of economic theory, the new one is the practice of randomized trials, as in biology and medicine. Random selection of students who are to participate in a policy intervention allows relatively simple measures of the difference between students belonging to the 'treatment group' (who benefited from the intervention) and the 'control group' (who did not) to yield unbiased estimates of the intervention's impact.<sup>93</sup>

Randomized experiments belong to the field of policy evaluation as much as to economics.<sup>94</sup> The evaluation of many types of policy intervention through randomized experiments, long practiced in the USA, is becoming a standard tool in development programmes, notably World Bank-funded projects. Such experiments remain less frequent in the field of education, but have already provided some credible evidence about the effectiveness and cost-efficiency of alternative policy interventions. Experiments are not without their weaknesses, however. They are usually small scale, so the generality of their results and the possibility of scaling up are not guaranteed.<sup>95</sup> Where benefits are widely expected, they may induce selective migration, and other practical difficulties can also result in the selection of students 'for treatment' not being fully random. Furthermore, conducting experiments in education may be difficult on ethical, political or financial grounds, e.g. it can be difficult to justify the exclusion of some people from a potentially beneficial intervention.

The evaluation of many types of policy intervention through randomized experiments is becoming a standard tool in development programmes

93. Todd and Wolpin (2003) demonstrate that structural estimates of education production function parameters and experimental estimates of policy intervention effects are not directly comparable. The former reflect a technical relationship between inputs and outputs and answer such questions as 'How would a change in class size, holding all other inputs constant, affect achievement?' The latter measure the total direct and indirect effects of an intervention, answering such questions as 'What would be the total effect of a change in class size on achievement, that is, not holding other inputs constant?' This difference, and the fact that structuralists and 'experimentalists' tend to form separate scientific communities, justifies their separate consideration.

94. See Burtless (1995), Dufo and Kremer (2003) and Newman, Rawlings and Gertler (1994) for further discussion of this point.

95. On this, see Dufo (2003).

### Children belonging to disadvantaged social backgrounds benefit most from reduced class size

One way to generate experimental data without raising such issues is to randomize the phasing in of programmes that are in any case to be gradually introduced.

Experimental estimation techniques may find another application in non-experimental settings where history or institutions have caused variation in education policies. Such circumstances are termed 'natural experiments' or 'quasi-experiments'.

Better evidence as to the efficacy of resource improvements to schooling is emerging from recent country studies using the new approach. Where key differences in schooling conditions, as between different groups of pupils, have been sharply controlled, stronger contrasts in inputs and outcomes can be seen. Results from some 'random' and 'natural' experiments, briefly discussed below, suggest, somewhat more clearly than production function evidence, that well-designed resource policies can have a strong qualitative impact on the outcomes of schooling.

#### *Randomized experiments*

Experience from industrialized countries is available from as early as the late 1980s. The Tennessee Student-Teacher Achievement Ratio (STAR) project was a large-scale randomized experiment on the impact of class size on pupil achievement. It followed almost 12,000 pupils from kindergarten to grade 3. The chosen schools were large, with at least three classes per grade. Pupils were randomly divided between a treatment group with small classes of thirteen to seventeen students and a control group with regular classes of twenty-two to twenty-five students, with or without a full-time teacher's aide. Standardized tests were administered to all students at the end of each year. Estimates of the project's impact were made on the basis of differences in test scores, within each of the seventy-nine participating schools, between pupils belonging to the treatment and control groups.

The results show that pupils in small classes perform better than those in regular classes and than those in regular classes with aides. The difference is established in the kindergarten class at about five percentage points. It widens to 8.6 points in the first grade, and stays at five

or six points in the second and third grades (Krueger, 1999: 511). Thus, most of the impact is observed after the first year, with the additional impact of subsequent years still being positive, but smaller. Finally, the impact tends to be larger for students who were eligible for free lunches and for black students and inner-city students, which suggests that children belonging to disadvantaged social backgrounds benefit most from reduced class size.

These differences appear to have continued throughout the students' subsequent school careers. The test scores of black students included in the initial study increased by seven to ten percentage points from kindergarten to grade 3, as against three to four percentage points for white students. Test scores in further grades, after the end of the experiment, increased by five points per year for black students and 1.5 points for white students. The probability of the black students taking entrance examinations for college was also increased much more substantially than that of the white students (Krueger and Whitmore, 2002).

The results of STAR provide robust evidence that class size affects achievement in the context of a high income country. It shows that reducing class size in the early school years appears to have long-run effects, especially in terms of reducing social inequalities in test scores and access to higher education. If this were generally true, developing preschools and improving primary schools could be expected to have larger returns than targeted cognitive-skill-based policies for adolescents (Carneiro and Heckman, 2003).<sup>96</sup>

In the last few years, studies related to developing countries have also begun to appear. In Kenya, the impact of school meals in a sample of twenty-five treatment and twenty-five control preschools was examined (Vermeersch, 2002). It was found that attendance increased by 30%. The programme was also expected to have an impact on test scores, but because organization of the meals cut into instruction time, an increase in test scores (by 0.4 standard deviation) was observed only in schools in which teachers had received substantial training. A separate study assessed the impact of provision by an NGO of free uniforms, textbooks and classroom construction, using fourteen poorly performing

96. The benefits of early childhood intervention in the United States are the focus of a sizeable literature that includes randomized experiments. See Currie (2001) for a survey.

Kenyan schools, half being the control group (Kremer et al., 1997). The intervention reduced dropout rates in treated schools, resulting over five years in a 15% increase in the average amount of schooling completed. Furthermore, many parents transferred their children from nearby schools to treatment schools, leading to an increase in class size by 50%, without deterioration in test scores. Thus, parents were willing to accept larger class sizes in return for lower direct costs as long as the increase in size did not have a major impact on achievement. This suggests that, in some contexts, a strategy of increasing class size so as to spend more on improving school inputs and reducing fees would be beneficial.

Again in Kenya, it was found that the provision of textbooks to schools increased test scores by about 0.2 standard deviation. The impact was concentrated among students who had scored in the top two quintiles in tests administered before the programme was started, however, and there was no impact on the other 60% of pupils. The intervention tended to be biased towards children belonging to relatively privileged families (Glewwe, Kremer and Moulin, 2002).<sup>97</sup>

In a randomised experiment conducted in the Philippines, thirty schools were allocated either to a control group or to one of four treatment groups, which received school meals, pedagogical materials for teachers, or structured meetings between parents and school officials combined with one of the other interventions. The provision of pedagogical materials had a considerable impact on dropout rates after one year, and all interventions significantly affected test scores (Tan, Lane and Lassibille, 1999).

A range of randomized experiments has been conducted in India. Among the studies was one showing that appointing a second teacher to single-teacher schools increased females' school participation but had no impact on test scores. A programme that provided teacher training and other inputs to preschool teachers who had begun teaching after receiving only minimal training was also found to have little impact on test scores (Banerjee and Kremer, 2002). However, a large and innovative experiment begun during the 1990s in urban Indian schools demonstrates the positive impact on learning of reducing class size, targeting

attention to ability groupings and, especially, providing child-friendly teaching. In this programme, lagging pupils are taken out of school for two hours per day to be taught by educated young local women who have undergone a short training programme. Existing classes are thus reduced in size for a time, for both low- and high-achieving pupils. The former group also benefits from more focused attention on their difficulties and the latter benefit from higher average abilities among their peers. As Box 2.4 reports, the test scores of both groups increased significantly – but particularly so for those in the lower-achieving category. It is likely that a crucial ingredient for success was the selection of local women who were strongly motivated to work with the children in the schools concerned. The experiment nonetheless demonstrates both the potential impact of class size on achievement and the potential effectiveness of using lower-cost teaching solutions for targeted purposes.

#### *Natural experiments*

Studies involving natural experiments are based on similar techniques to those used in studies based on randomized experiments: in both cases, a specific source of variation in a potential determinant of achievement can be identified and studied. Natural experiments in education are usually the result of the application of some policy rule that is independent of the usual functioning of schools. Their obvious disadvantage relative to randomized experiments is that researchers do not have control over them, and data may be lacking. Thus, the number of exploitable natural experiments may not be very high.

The case of Israel affords an opportunity to investigate the impact of class size on pupil performance. Israeli schools are allocated teachers according to a rule formulated by the medieval scholar Maimonides, which says that class size should not exceed forty pupils. As a result, a school with, say, thirty-nine pupils will have one teacher, but a school with forty-one pupils will have two. This generates a systematic, predictable relationship between the number of pupils enrolled in a school and the PTR that is uncorrelated with unobserved characteristics of pupils that affect learning.<sup>98</sup> When comparing class-size data with test scores for 2,000 classes of third- to fifth-grade pupils, a significant

**The provision of pedagogical materials had a considerable impact on dropout rates after one year**

97. The authors note that Kenyan textbooks are written in English and reflect a curriculum designed for elite families in Nairobi, so they may be difficult for rural children to grapple with.

98. Such a correlation could still exist if application of the rule differed by neighbourhoods' socio-economic makeup – for example if privileged areas more easily obtained additional teachers as the PTR neared 40:1. Angrist and Lavy (1999), however, provide evidence that the rule is quite strictly observed.

### Box 2.4 The impact of a remedial education programme in India's urban slums

In the slums of a number of Indian cities, Pratham, an NGO, runs a remedial education programme in primary schools. Pratham hires young women (*balsakhi*), holding the equivalent of a high school degree, who belong to the same community as the children. They teach one group of fifteen to twenty lagging children, who are withdrawn from their normal classes for two hours every morning, and another group in the afternoon. Teachers receive two weeks of initial training and further training during the year. The *balsakhis* focus on basic numeracy and literacy, which children should have learned in the second and third grades. The expected benefits of the intervention include the impact of remedial education on the children who attend the special classes and the impact of reduced class size for all.

The programme was used to conduct a randomized experiment in the cities of Mumbai, where the *balsakhi* programme started in 1994, and Vadodara, where it started in 1999. In Vadodara, the programme was extended over 2000-02 to the ninety-eight eligible schools not yet covered. Half were randomly selected to receive *balsakhis*, and their results were compared with those of the half that did not receive the intervention. A similar procedure was followed in Mumbai. Literacy and numeracy tests were administered in all schools before the

intervention began and at the end of each school year. More than 15,000 students were included, which allowed a reliable estimation of its impact.

The results revealed a significant impact on achievement, which was remarkably similar in both cities. Test scores of children who benefited from the programme improved by 0.12 to 0.16 standard deviation in the first year and 0.15 to 0.30 in the second year. The results were even stronger for the lowest achievers. Children in the bottom third of the distribution improved their test scores by 0.22 standard deviation in the first year and 0.58 in the second. Moreover, the results suggest that the direct impact of *balsakhi* sessions on children who attended them accounts for most of this improvement, rather than the impact of reduced class size for the other children. This implies that *balsakhis* were more effective in their separate classes than regular teachers were in the combined classes. The study concludes that both the efficacy of the remedial education programme and, more generally, the feasibility of strongly improving test scores at very low cost have been demonstrated. The estimates suggest that reducing class size by hiring a *balsakhi* is at least twice as effective as reducing class size by keeping children with regular teachers, even though *balsakhis* are paid only a fraction of teachers' salaries.

Source: Banerjee et al. (2003)

99. Angrist and Lavy (1999) discuss the results, which are subject to debate. Similar studies sometimes produce conflicting results. In the state of Connecticut, United States, where similar maximum class-size rules allow comparison, no impact on test scores was found for variations ranging from ten to thirty pupils (Hoxby, 2000). Comparing primary schools in Connecticut and Israel is not straightforward, however, if only because class size is much larger in average Israeli schools.

100. The authors state that interviews with officials then in charge revealed that there were no consistent rules for allocating resources to schools for black children. Hence, they argue, the wide variations in PTR observed between those schools can be considered random, and exogenous with respect to family characteristics affecting learning. The natural experiment here is, somewhat paradoxically, the absence of a well-defined school policy due to neglect of the well-being of the black population under apartheid.

101. The impact on mathematics scores was not significant.

negative impact of class size on the reading and mathematics scores of fifth-graders is found: a decrease in class size by a standard deviation (6.5 pupils) results in an increase in reading test scores by 0.2 to 0.5 standard deviation and an increase in mathematics test scores by 0.1 to 0.3 standard deviation. The effects on the achievement of third- and fourth-graders, however, were often insignificant.<sup>99</sup>

As regards developing countries, the notion that class size is important for cognitive achievement is further strengthened by a study of South Africa at the end of the apartheid regime (Case and Deaton, 1999). During the apartheid period, blacks had no political representation and therefore no control over government funding of education, which was heavily centralized. Furthermore, controls were imposed on their

place of residence, preventing internal migration, except under tightly defined circumstances. Consequently, parents of black children had no influence on the quality of the schools their children attended.<sup>100</sup> Data collected in 1993 revealed a significant impact of PTR variation on years of schooling completed, on current enrolment status and on test scores of black children. For example, decreasing the PTR from forty (the mean value for black schools) to twenty (the mean for white schools) increased schooling by 1.5 to 2.5 years and reading test scores by the equivalent of what was usually delivered by two additional years of schooling.<sup>101</sup>

This approach has also been used, with data from the 1996 South African census, to estimate the impact of school quality. Significant negative effects of the PTR on educational attainment, on

the probability of employment and on the returns to education are revealed. For example, reducing the PTR by five students would increase the return to education by 1% (Case and Yogo, 1999).

Other research using these methods addressed the effectiveness of in-service teacher training (Angrist and Lavy, 2001). As with class size, production function evidence on teacher training is mixed, but it tends to focus on initial training, while in-service training could plausibly have a stronger impact. A few Jerusalem schools received funds earmarked for in-service teacher training in 1995, and the study matched these schools with comparable schools in the same area that did not receive such funds.<sup>102</sup> It found that teacher training improved test scores by 0.2 to 0.4 standard deviation in secular schools. The training programme focused on pedagogy rather than subject content, and was inexpensive compared to class-size reduction.

In summary, it appears that this newer evidence deriving from random or natural experiments can indeed reveal the impact of particular policy changes on learning achievement, at least in particular contexts. These studies do not rely on theoretical modelling, so a wider range of interventions has already been considered than in the education production function literature. Empirical modelling is made simpler by the identification of a particular, externally determined source of variation in school resources, which is a matter of policy analysis, not of statistics or econometrics. The dissemination of the results beyond academic circles is thus bound to be easier.

The application of these methods is as yet in its infancy. Yet it is already clear that strong evidence of the importance of school resources in affecting learning outcomes has been demonstrated in industrialized countries and, even more so, in developing countries. The case here is considerably stronger than that provided by the production function literature for the notion that smaller class size, more textbooks, teacher in-service training and teaching strategies that focus upon the needs of the learner are resource strategies that make a difference, particularly in poorer countries and particularly for less able groups.

## The role of the organization and social context of schools

Teaching and learning in the classroom do not take place in isolation from the functioning of schools as organizations, nor from their social context. Organizational weaknesses of schools are increasingly being pointed to as a major cause of low learning achievement, especially in the case of government schools in developing countries. For example, a landmark field study of rural primary schools in northern India (PROBE Team, 1999) provided striking evidence of a lack of teacher motivation (e.g. absenteeism and a tendency to supervise rather than actually teach) but also noted the difficult working conditions teachers face and the lack of control and support from education authorities. Meanwhile, education expansion, whether at primary level in developing countries or at secondary and higher levels in industrialized countries, results in changes in the social composition of the student population. The functioning of schools can be affected by this in ways that are quite separate from the impact of resources or teaching practices.

This section examines recent evidence concerning the functioning of schools as organizations and their interaction with their social context. One approach proposes radical changes in the structure of teacher incentives, arguing that this could have a more effective impact on learning outcomes than would further rises in expenditure or changes in teaching practices. Increasingly popular among economists and policy makers, such 'incentive-based' policies have been implemented in industrialized and developing countries, so there is growing evidence concerning their benefits and shortcomings. A second approach, maintaining that neither resource levels nor incentive structures can properly account for differences in schooling outcomes, gives weight to political and sociological factors.

## Promises and pitfalls of incentive-based policies

While research on school and teacher effectiveness has addressed the behaviour of teachers as pedagogues, economists have investigated their behaviour as employees of schools or of education authorities. This work has typically examined the extent to which a

**Organizational weaknesses of schools are increasingly being pointed to as a major cause of low learning achievement**

<sup>102</sup> 'Matching' is an econometric technique that does not explicitly rely on the experimental paradigm but similarly allows a reduction in endogenous biases. While it has become extremely common in labour economics, its use in the economics of education is still infrequent, as Glewwe (2002) notes.

**'Merit pay' systems tested in the United States do not seem to have been successful**

teacher's 'productivity' in the classroom is linked with remuneration and whether reinforcing any such link can improve learning outcomes.<sup>103</sup> There is some evidence that this link is, at present, weak. For example, in urban state-run secondary schools of northern India, teacher characteristics that determine pay, such as the quantity of education, training and experience, have little impact on pupil achievement, while the quality of teacher education, represented by cognitive skills, affects achievement but not pay (Kingdon, 1996). By contrast, in private schools a direct relationship between pupil achievement and teacher pay appears to at least partly explain why they are more efficient than state schools (Kingdon and Teal, 2003).

'Merit pay' systems tested in the United States do not seem to have been successful. Teachers generally responded by increasing the quantity rather than the quality of teaching, and uniform pay structures evolved, defeating the purpose of the reforms.<sup>104</sup> On the other hand, a programme introduced in Israeli high schools, whereby teachers received cash bonuses for improvements in their students' performance on matriculation examinations, had a significant impact on English and mathematics results (Lavy, 2003).

Merit pay nevertheless seems quite a narrow approach to teacher incentives. First, it can have perverse side-effects. For example, in a randomized experiment conducted in Kenya, prizes were attributed by parent-run school committees to teachers whose pupils had low dropout rates and performed well on exams. Teachers responded by manipulating exam results rather than teaching better, and there was little impact on real achievement (Glewwe, Nauman and Kremer, 2003). Second, where salaries are relatively high, additional pay may not be a major motivating factor for teachers, given the specificity of their profession.<sup>105</sup> Recent research in public schools in the United States suggests that, while institutional arrangements do result in a weak relationship between teacher salaries and performance, student characteristics rather than salary differentials explain teachers' decisions to change schools.<sup>106</sup> Meanwhile, recent evidence from Mexico's Carrera Magisterial programme indicates that providing additional training while raising salaries and improving school resources can

increase pupil achievement. The impact of that programme appeared greatest when training was aimed at increasing teachers' practical experience and developing content-specific knowledge, and when supervision by school administrators was high (Lopez-Acevedo, 2004). This finding suggests that policies focusing on teaching practices, resources and incentives may be complements rather than substitutes.

Similar remarks apply to decentralization and privatization, which have been advocated as giving parents more control over the functioning of schools and over teacher behaviour. There is indeed evidence that increasing local financing can improve learning, and several studies have found private schools to be more efficient than public ones.<sup>107</sup> But any general appraisal of such radical changes in school system organization involves considering other outcomes of schooling than cognitive achievement, given the much broader socio-political impact of such reforms. In Kenya, decentralization is reported to have created incentives for local communities to build too many small schools, to spend too much on teachers relative to other inputs and to set school fees at levels that prevent many children from attending (Kremer, Moulin and Namunyu, 2003). While the case for increasing teacher and school accountability to parents is simple enough, designing reforms to this end appears to be a difficult task.

### Schools as social institutions

Explaining how and why changes in resource allocation or teaching practices have – or fail to have – an impact on cognitive achievement requires understanding how teachers and students interact, not only as teachers and learners but also as persons in the social context of the school. Education is not only a set of outcomes, it is also a process. Given the amount of time spent in it, the behaviour of teachers and students is bound to be determined as much by the quality of this process as by its anticipated returns. Improving learning outcomes thus does not only involve providing sufficient resources for efficient use by teachers who use best-practice methods, with an adequate incentive structure. It can also involve affecting the way students and teachers experience life and work at school throughout its duration.

103. Much of this literature assumes that schools *should* be likened to achievement-maximizing private firms in a competitive market, and is thus not neutral ideologically. Whatever one's opinion about this assumption, the gathering of evidence on implementation of the increasingly frequent reforms using this approach remains important.

104. The evidence is surveyed by Hanushek (2002b).

105. This, of course, is far from being the case in many developing countries. In recent years, salary levels have often sunk so low that the very subsistence of teachers and their families is threatened. Under these circumstances the link between pay and performance is difficult to deny. See the discussion on teacher earnings in Chapter 4.

106. See Hanushek and Rivkin (2003), Hanushek, Kain and Rivkin (1999), Hanushek, Kain and Rivkin (2004) and Jepsen and Rivkin (2002).

107. On both points, see Glewwe (2002).

Studies in the sociology of education suggest, first, that the quality of learning depends on student behaviour more than on anything else, and that student behaviour responds directly to socialization processes that take place at school.<sup>108</sup> Ethnographic and sociological studies of high schools in the United States have shown that students divide into distinct social groups, each associated with an ideal behaviour to which its members strive to conform. Ideals include the attitude towards the school in general and towards learning in particular. Student effort is influenced by the social group to which a student belongs and by the ideal of that social group. Second, schools are not neutral vis-à-vis student behaviour. They try to impart not just cognitive skills but also a series of psychological and behavioural traits that define their own social ideal. Students whose family and peer group have ideals close to that promoted by their school will tend to demonstrate greater learning effort and achieve higher levels of cognitive skills. Others will try to escape the contradiction between their own and their school's ideals by rebelling – reducing learning effort, among other things.<sup>109</sup> Thus, an important way for schools to improve learning achievement is to reduce the social distance between their own social ideal and that of their students, so that fewer students choose to belong to social groups whose attitude clashes with learning.<sup>110</sup>

This conceptual framework may shed new light on the situation in many developing countries. The education policies pursued after independence in many states often explicitly promoted a social ideal, whether expressed in terms of national unity or a more explicitly socialist or other ideology. Initially this concerned only the parts of the population that happened to be able to attend school, whether groups with elite status or ethnic or other favoured social groups. Expansion of the school system beyond those groups raised crucial issues concerning the content of teaching. Meanwhile, the ideological wave that had accompanied decolonization receded, and resources also became scarcer during the 1980s. Today, an array of factors – continued resource scarcity, discrepancies between teachers' and students' social backgrounds, clashes between the social content of the curriculum and students' family backgrounds, the lack of an explicit social goal for education policy – may help explain the

widespread inefficiency of state school systems in developing countries in a subtler and more relevant way than single-factor explanations such as inadequacy of teacher pay.

## Conclusion: what we know about what matters for education quality

There is ample evidence that the quantity of education a person receives (measured as the number of years spent in the school system) goes hand in hand with the quality of that education (usually somewhat narrowly defined as cognitive skills but in fact including non-cognitive skills, values and other psychological and behavioural traits acquired through schooling). The latter aspect has intrinsic value and is also associated with many and various private and social returns. These are not limited to income but also include advantages derived from a range of market and non-market activities.

International assessments of cognitive skills suggest that school quality differs widely among and within countries. In particular, children who live in developing countries not only receive fewer years of education but also reach lower achievement levels. Meanwhile, though the evolution of test scores over the years is difficult to assess and interpret, clearly their stagnation in developed countries in recent decades represents an important puzzle.

Identifying the determinants of better learning outcomes so as to produce policy-relevant conclusions is an arduous task that requires using approaches from different social sciences. The learning process is extremely complex. It first and foremost involves relationships between teachers and students following a given curriculum and teaching practices, but it also takes place in a broader social context. These relationships are further conditioned by the resources available to schools, the incentive structure teachers face as employees of schools or education authorities, and the correspondence between the values promoted by schools as social institutions and those that prevail in students' families and society at large.

## The stagnation of test scores in developed countries in recent decades represents an important puzzle

108. This approach has been explored recently by Akertof and Kranton (2002).

109. A similar representation could be given of teacher behaviour.

110. This concept can be applied to the history of secondary education in a number of industrialized countries. Akertof and Kranton (2002) show that the democratization of school systems in the last third of the twentieth century led to growing clashes between earlier ideologies – that of the French republican bourgeois elite, for instance – and those of the more socially diverse population that was increasingly entering schools.

**Designing more adequate incentive structures for teachers appears a natural complement to resource policies**

This complexity is reflected in the findings of this chapter.

- First, an examination of country-specific experience reveals that very different school systems have produced either low or high average levels of achievement. Much is to be learned from further comparisons among and within countries.
- Second, debates within the education community regarding adequate teaching practices are not settled. Hence, experimenting with alternative practices is necessary, especially as progress towards Education for All results in enrolment of children from communities that were formerly socially distant from the school system.
- Third, while the education production function literature has yielded no easy consensus on the impact of specific resource policies on achievement, there is evidence, especially from recent experimental studies, that sufficient resources are necessary if education of acceptable quality is to be attained, and that well-implemented increases in resources are an important means of improving educational quality in developing countries.
- Fourth, education policies need to address the efficiency of resource use in schools. Designing more adequate incentive structures for teachers appears a natural complement to resource policies. However, the limited evidence on ideas such as merit pay, decentralization and privatization show that they can have perverse side effects that either reduce their impact on cognitive skills or affect other educational outcomes.

In the end, improving learning outcomes in schools around the world is not only a matter of implementing a set of adequately designed technical measures. That the experimental evaluation of policy interventions is becoming a standard tool in policy design is certainly a welcome development, but it does not solve the question of which interventions are tested or what their goals are. Schools are social institutions in which day-to-day educational processes interact with the shaping of educational outcomes. The cognitive and non-cognitive outcomes of this process may sometimes conflict. As a larger proportion of children are socialized for an increasingly extended part of their childhood, adolescence and early adulthood, questions of curriculum content and of contrast between student background and aspiration present new challenges for the quality of schooling that cannot be addressed only by technical means. The politics of the process, as well as the details of its resourcing and pedagogy, have become increasingly important to its solution. ■