QUESTIONNAIRE FRAMEWORK FOR PISA 2003

March 2003

Project Consortium:

Australian Council For Educational Research (ACER)

Netherlands National Institute for Educational Measurement (CITO group)

Educational Testing Service (ETS)

National Institute for Educational Policy Research (NIER, Japan)

Westat

TABLE OF CONTENTS

POSSIBLE THEMES FOR PISA 2003 THEMATIC REPORTS

INTRODUCTION

1. PISA seeks to provide policy-relevant information on student outcomes through the assessment of the achievement of 15-year-olds in reading, science, mathematics, and aspects of cross-curricular competence. In addition, through the collection of information on students and their schools PISA allows to identify social, cultural, economic and educational factors that are associated with student performance. Using the data from these two questionnaires, analyses linking context and information with student outcomes allow PISA to address differences:

- between countries in the relationships between student level factors (such as gender and social background) and outcomes;
- in the relationships between school level factors and outcomes across countries;
- in the proportion of variation in outcomes between (rather than within) schools, and differences in this value across countries;
- between countries in the extent to which schools moderate or increase the effects of individual-level student factors and student outcomes;
- in education systems and national context that are related to differences in student outcomes across countries; and,
- changes in any or all of these relationships over time.

2. In its call for tender for PISA 2003, the BPC established the main policy issues it sought to address with PISA 2003. In particular, the BPC required PISA 2003 to portray important aspects of learning and instruction in the field of mathematics, including the impact of learning and teaching strategies on achievement as well as the impact of school organisation and structures in promoting active student engagement with learning. The BPC also required PISA 2003 to address issues related to: mathematics efficacy and students' engagement, motivation and confidence with mathematics; mathematics and gender; and students' planned educational pathways. Finally, the quality of the school's human and material resources; issues of public/private control, management and funding; school level information on the instructional context and institutional structures were considered important issues for PISA 2003.

3. This document seeks to operationalise these goals within the design constraints and instruments that are underlying PISA 2003. This is accomplished through the following steps:

- First, an organising framework is established that allows to map the BPC's policy issues for PISA 2003 against the design and instrumentation of PISA. The objective is to facilitate choosing themes that combine policy relevance effectively with the strengths of PISA design.
- Having identified a conceptual structure from which relevant themes can be established, a set of criteria are developed for defining and operationalising the BPC's policy priorities within this conceptual structure.
- Third, proposals for thematic reports for PISA 2003 are outlined, with each proposal presenting a brief review of relevant literature, the specific policy questions the report could address, and how the theme would be operationalised in the PISA 2003 context questionnaires.

Organising framework

4. To facilitate a systematic approach to the organisation and prioritisation of themes, the framework for the OECD education indicators (INES) is applied which organises policy issues that might be considered in PISA 2003 by two dimensions:

- the level of the education system to which the resulting indicators relate; and
- whether they relate to outcomes or outputs, to policy-amenable determinants of these outcomes or outputs, or to given constraints at the respective level of the education system.

5. The framework considers four *levels* that relate both to the entities from which data might be collected and to a recognition that national education systems are multi-levelled. The four levels are:

- the education system as a whole,
- the educational institutions and providers of educational services,
- the instructional setting and the learning environment within the institutions,
- the individual participants in learning activities.

6. A differentiation between levels is not only important with regard to the collection of information, but also because many features of the education system play out quite differently at different levels of the system. For example, at the level of the students within a classroom, the relationship between student achievement and class size may be negative, if students in small classes benefit from improved contact with teachers. At the class or school level, however, students are often intentionally grouped such that weaker or disadvantaged students are placed in smaller classes so that they receive more individual attention. At the school level, therefore, the observed relationship between class size and student achievement is often positive (suggesting that students in larger classes perform better than students in smaller classes). At higher aggregated levels of education systems, the relationship between student achievement and class size is further confounded, e.g., by the socio-economic intake of schools or by factors relating to the learning culture in different countries. Past analyses, which have relied on macro-level data alone, have therefore sometimes led to misleading conclusions.

7. The second dimension in the organising framework groups the indicators at each of the above levels (i.e. the system, institutional, classroom and individual level) further:

- Indicators on observed outputs of education systems, as well as indicators related to the impact of knowledge and skills for individuals, societies and economies, are grouped under the sub-heading *output and outcomes of education and learning.*
- The sub-heading *policy levers and contexts* groups activities seeking information on the policy levers or circumstances that shape the outputs and outcomes at each level.
- These policy levers and contexts typically have *antecedents* – factors that define or constrain policy. The sub-heading *antecedents and constraints* represent these. It should be noted that the antecedents or constraints are usually specific for a given level of the education system and that antecedents at a lower level of the system may well be policy levers at a higher level (e.g. for teachers and students in a school teacher qualifications are a given constraint while, at the level of the education system professional development of teachers is a key policy lever).

8. The four levels and the three aspects can be visualised in a two-dimensional matrix.

	Column 1: Outputs and outcomes of education and learning	Column 2: Policy levers and contexts	Column 3: Antecedents and constraints
Individual participants in education and learning	***Cell 1:*** *Individual outcomes* For example: a) Reading, mathematics and science literacy b) Habits in relation to content domains c) Affective outcomes (e.g. attitudes to mathematics) d) Life skills *e)* Learning Strategies	***Cell 5:*** *Policy levers and contexts relating to individuals* For example: f) Engagement and attitudes to school life g) Confidence in mathematics h) Attendance i) Time on task j) Use of school resources k) Participation in extension, remedial and external programmes *l)* Homework behaviours.	***Cell 9:*** *Antecedents and constraints at the level of individuals* For example: m) Wealth (economic capital) n) Cultural capital, o) Social capital, p) Socio-economic status q) Parental educational level, r) Educational resources at home, s) Family structure, t) Ethnicity, u) Parental expectations and aspirations, *v)* Age and Gender.
Instructional settings	***Cell 2:*** *Outputs and outcomes at the level of classrooms/instructional settings*	***Cell 6:*** *Policy levers and contexts at the level of classrooms/instructional settings* For example: w) Class conditions and processes such as orderliness, teacher-student relations, teacher supportiveness x) Practices including homework, dealing with individual differences, lesson structures *y)* Course offerings, content of lessons and teacher achievement expectations.	***Cell 10:*** *Antecedents and constraints at the level of classrooms* For example: z) Teacher characteristics including age, experience, commitment, subject matter orientation, and pedagogical beliefs. *aa)* Student class characteristics
Education service providers	***Cell 3:*** *Outputs and outcomes at the level of institutions* For example: *bb)* Institution level aggregates of: Reading, mathematics and science literacy; Habits in relation to content domains; Affective outcomes (e.g. attitudes to mathematics); Life skills ; and Learning Strategies *cc)* Differences in outcomes for students of various backgrounds	***Cell 7:*** *Policy levers and contexts at the level of institutions* For example: dd) Instructional support including both material and human resources, ee) Policies and practices, including assessment, academic orientation, admittance policies, timetabling policies and *ff)* Climate including orderliness of the school, teacher morale, student-teacher relations, and leadership.	***Cell 11:*** *Antecedents and constraints at the level of institutions* For example: gg) The wealth, values and involvement of parents and the community in the school. *hh)* The type of school, its source of funding, its location and size.
The education system as a whole	***Cell 4:*** *Outcomes at the level of the education system* For example: *ii)* System level aggregates of: Reading, mathematics and science literacy; Habits in relation to content domains; Affective outcomes (e.g. attitudes to mathematics); Life skills ; and Learning Strategies *jj)* Equity related outcomes	***Cell 8:*** *Policy levers and contexts at the national level* For example: kk) Organisation of education (tracking, funding, school structures, length of the school year, day and week, and locus of control in the school) ll) Teacher qualifications and training requirements *mm)* School entry age, retention.	***Cell 12:*** *Macro-economic and demographic context* For example: • Gross Domestic Product • Distribution of wealth nn) Cultural homogeneity

9. While this mapping is useful for planning the coverage of the PISA questionnaires it is also important to supplement it with recognition of the dynamic elements of the education system. All of the cells in the framework are linked, both directly and indirectly and a range of important indicators that deal with the relations between the cells will be central to the outcomes of PISA 2003.

10. For example, consider an indicator that was central to the outcomes of PISA 2000: *What is the relationship between student outcomes and student background factors such as socio-economic status?* This indicator is directly concerned with the relationship between Cell

9 and Cell 1, and at the same time its further exploration is concerned with how data relating to cells 5 to 8 might influence this relationship.

Criteria for defining and operationalising the BPC's policy priorities

11. It is not only important to identifying possible policy-relevant themes but also to choose wisely from among the many possibilities so that the strengths of the PISA design are capitalised upon and PISA's contributions to policy-makers' and educators' needs is maximised.

12. This document uses a set of criteria that have been established by the INES Network A for defining and operationalising the BPC's policy priorities and for developing relevant thematic reports:

- First, they must be of enduring policy relevance and interest. That is, a theme should have *policy relevance*, capture policy-makers' attention, address their needs for data about the performance of their educational systems, be timely, and focus on what improves or explains the outcomes of education. Further a theme should be of *interest to the public*, since it is this public to which educators and policy-makers are accountable.
- Second, the themes must provide an internationally comparative perspective and promise significant added value to what can be accomplished through national evaluation and analysis. This implies that themes need to be both relevant (i.e., of importance) and valid (i.e., of similar meaning) across countries.
- Third, there must be some consistency in approach and themes with PISA 2000.
- Fourth, it must be technically feasible and appropriate to address the issues within the context of the PISA design. That is, the collection of data about a subject must be *technically feasible* in terms of methodological rigor and the time and costs (including opportunity costs) associated with data collection.

13. These criteria are explained in Annex 1 in more detail.

THEME 1: SCHOOL CHARACTERISTICS, ORGANISATION AND STRUCTURE

Elaboration

14. As an international comparative study, PISA has as its primary focus on differences among countries in students' mathematics, reading and science literacy. At the same time, however, it is important to recognise that most of the variation in student performance will be within countries. A range of research has shown that, in many countries, differences between schools account for a considerable part of this variation. PISA 2003 provides an opportunity to explore the variation between countries in school effects – that is the proportion of between student variance that can be accounted for by schools – and to explore some key variables that might cause variance between schools.

15. The design of PISA is such that those school characteristics that are *pervasive* or those that are related to important characteristics of the educational system are likely to be most worthwhile pursuing (see also Annex 1). These characteristics can be grouped into variables related primarily to the *structure of schooling* and those related to the *instructional context within schools*. It is proposed that the following eight aspects be covered in PISA 2003, with a focus on both describing the learning environment of students and identifying possible relationships between school characteristics and student outcomes.

Characteristics related to the structure of schooling:

- *Grouping of students*: Schools may differ in the study programmes they offer as well as with regard to the extent of ability grouping between or within classes.
- *Segregation of schools*: In many countries schools differ in their social composition (for example, socio-economic status, language background, gender).

- *Management and Financing*: Schools may differ with respect to their management (for example the degree which schools have autonomy in their decision making) and financial structures (for example public/private).
- *School Resources*: The extent to which school differ with regard to their equipment with human and material resources.
- *School size and location*: The extent to which schools differ with respect to their location and size.

Characteristics related to the instructional context within schools:

- *Learning Time*: Schools have varying degrees of flexibility in the allocation of learning time of students.
- *Student support policies***:** Schools may differ in their support policies and practices, especially for low- and high achieving students.
- *School and Classroom Climate*: Learning outcomes in schools may be influenced by the general climate at both school and classroom level.

16. These variables are, of course, often inter-related. Having a strongly centralised educational system, for example, may have consequences for the instructional context variables, for example, less flexibility in the allocation of learning time and support policies. Furthermore, there is a body of research that shows that highly tracked educational systems may increase segregation since socio-economic status influences student (or family) choices of educational tracks. In such systems schools may prove to be agencies for reproducing social stratification (Douglas, 1964; Heather, 1969; Kershaw, 1992; Persell, 1992).

17. Research on the effects of school characteristics on performance dates back (at least) to the sixties: Coleman *et al.* (1966) and Jencks (1972) concluded that schools were not major determinants of a child's achievement, particularly when contrasted with the influences of family background on student outcomes. Later studies, although providing similar evidence, valued the facts quite differently: schools did matter, to some degree. The common view now is that since schools and teachers are the most obvious factors in the learning endeavour that are amenable to policy intervention, potential effectiveness enhancing factors should be of central interest to educational research studies (see an overview in Teddlie and Reynolds, 2000).

18. Clearly the importance of school effects is an area of important ongoing debate. Social intake and school characteristics are often highly correlated, which makes it difficult (particularly in cross-sectional studies) to determine what proportion of any observed school differences might be *true* school effects (in the value-added sense) and what proportion might be due to the social context of a school. Mayeske *et al.* (1972) in their re-analysis of the data used in the Coleman report of 1966 assumed that the joint variance between school and background variables should be considered as part of the school, not of the background effect.

19. Though variables describing global characteristics of schools are often considered as distal factors and have low correlations with students' outcomes, they play an important role in shaping the context in which each individual classroom or teacher is placed, thus favouring or precluding the successful implementation of a range of instructional practices. That is they set a context that may or may not support activities that are known to have more direct relationships with achievement and student willingness to invest in learning activities. For example, most of the literature on innovation confirms that when new instructional practices are implemented, lack of school-wide organisational support or problems with traditional school structures are a frequent cause of failure in achieving expected or sustained improvements (for example, Stringfield *et. al.*, 1996; Conti *et. al.* 2000; Lee, 2001).

20. In this respect, organisational and structural aspects are particularly relevant to policy makers and may be important policy levers. While national or district authorities cannot easily monitor what happens in each specific class or lesson, school-level policies may appear as more malleable variables when large-scale reforms or innovative programmes are being implemented.

21. To the extent that many school global characteristics often show limited variance in single countries, but can vary across countries (as is often the case for variables such as class organisation, time scheduling, courses offered, teachers and principals' educational 'culture' or 'philosophy'), international studies like PISA can help identifying organisational factors that occur more frequently in successful school systems than in others.

22. Preceding a description of the range of variables and issues that might be explored in PISA, the following discusses in general terms some of the kinds of issues that data on school organisational policies and practices collected through the PISA 2003 student and school questionnaires can be used to inform.

- First, to what extent do the effects of organisational variables described as 'positive' in the literature actually appear as positive (in terms of student achievement and engagement in learning) in most or all OECD countries? In other words, to what extent are the policy recommendations that can be derived from research findings in certain school systems valid for other systems, other cultures and other societies? Which specific dimensions appear to have the most significant and internationally consistent effects? Are some of these policies effective in promoting achievement, but not motivation or strong engagement, or the reverse? Is there some evidence that, as suggested by the literature, when a range of strategies covering multiple aspects of the school context are combined in the same school setting, the combined effect is multiplicative rather than additive – thus pointing at the value of comprehensive, articulated, and 'school-wide' programmes rather than implementation of separate 'ad hoc' strategies? If so, which combinations or which configurations of school policies appear to be the most promising?
- Second, to what extent will the PISA results support that some of these policies may be effectively used to change important instructional dimensions at the classroom level? While the constraints on the PISA design do not allow for a full exploration of the direct effects of differences in school organisational policies on classroom instruction practices, some of these effects can be investigated indirectly, using the information collected through the PISA 2003 student and school questionnaires. For instance, are variations in the size of school (or of possible semi-autonomous units within the school) associated with differences in the student reports on teacher/students relations? Are school-level support policies echoed by positive student reports on the supportiveness of teachers? Do school-level strategies aimed at ensuring a productive and work-oriented environment or a positive school climate translate into low levels of classroom disruption?
- Third, how do the PISA countries compare in terms of the organisational policies used in their schools? Do any 'dominant' patterns emerge in certain countries or groups of countries, and if so, are these patterns associated with high levels of average achievement? To what extent can they be considered as an integral part of the educational 'culture' of the school system? For example, to what extent do comprehensive mainstreamed systems tend to favour student-support oriented and non-selective school strategies, compared to tracked systems? Are centralised school systems characterised by highly uniform organisational models? Do decentralised school systems appear to have a stronger potential for favouring a wide range of creative innovations in their schools' organisation? If so, is this positive characteristic of decentralised systems counterbalanced by possible disparities between schools with more effective or less effective policies?
- Fourth, since changes in certain aspects of school organisation are (together with changes in curriculum) one of the most frequent targets of district- or nation-wide policies, some of the indicators collected in the PISA 2003 instruments could be selected for use as trend indicators in future cycles. Time-series on such variables as admission and transfer strategies; on instructional time and time scheduling policies; on parental and community support practices; on interdisciplinary team-teaching (to quote only a few examples) could be of interest both for international comparisons and for monitoring possible school restructuring policies implemented at the national level.
- Fifth, equity issues are extremely important for policy-makers and educational segregation can be viewed as a factor increasing inequities in a society. This is of particular importance

in societies where schools differ substantially with regard to the social background of their students. Socio-economic background of the student body, is correlated with other school characteristics like school ownership, academic programmes, and quality of resources. To what extent is the effect of average SES on school performance enhanced or diminished by the homogeneous or heterogeneous grouping of students, and how is it related to the existence of a private school sector and selectivity of schools? Are there different impacts of urban and rural poverty on student performance?

School characteristics related to the structure of schooling

Grouping of students

23. The effect of *grouping of students* in different schools, different curricular tracks, and different ability groups is one of the oldest issues in educational science and of primary policy interest (Kulik and Kulik, 1982; Oakes, 1985; Postlethwaite, 1995). This issue is generally referred to as *integration* versus *differentiation*, or *homogeneous* versus *heterogeneous* grouping. Research has shown that educational systems can be roughly classified according to this dimension as integrated, differentiated, or mixed. A central question regarding variation in student performance is: *how much of the within country variability in student performance may be due to the specific grouping arrangements, either intended or coincidental*.[1]

24. In many countries, the policies related to this aspect (such as tracking, study programs offered, grade retention) are part of the structure of the educational system. However, school level policies can also play an important role in determining the extent to which the students receive their daily instruction in homogeneous or heterogeneous groups. Admission and transfer practices, decisions on grade retention, on establishing ability groups, on students' allocation to the various classes or various study programs may, for example, strongly reduce or increase the homogeneity of the group where each individual student is actually placed.

25. While heterogeneous grouping is often recommended on the basis of the existing literature (see above), the idea that homogeneous groups are best suited to address students' needs is still prevalent, in many schools, among both parents and school staff - even in school systems with mainly mainstreamed or untracked structures. This may result in the implementation, by certain schools, of 'overt' or 'covert' selection/tracking strategies, such as use of achievement tests when admitting students, transfer to other schools of students with behaviour problems, harsh requirements for promotion to the next grade, allocation of the most able students to specific classes on the basis of optional courses — such as foreign languages considered as more 'difficult' than others.

26. Most of these practices are likely to have detrimental effects on the achievement of students in low ability groups, on their engagement in learning activities, and on their aspirations for further educational attainment (Gamoran and Berends, 1987). In particular, controlling for other students' characteristics, grade retention is associated with higher rates of future dropout (Grissom and Shepard, 1989). On the other hand, however, heterogeneous settings without specific accommodations for low achievers may expose those students to unequal competition with more able students, and hence be detrimental for their motivation and self-concept.

27. Alternatively, schools can implement strategies enabling their teachers to 'work well' with highly heterogeneous classes, while avoiding frustration among disadvantaged or less able students. This may include a range of organisational and teaching practices, such as:

- using specialised teachers to help slow learners improve fundamental skills (for example, reading);

[1] This research question is also related to the theme 'Students educational pathways'.

- implementing 'incremental' assessment systems that provide each student with frequent recognition for meeting individually referenced short-term learning goals, rather than (or in addition to) more traditional assessment practices (Beady *et al.*, 1981);
- favouring multicultural education practices that make use of the strengths and socio-cultural experiences of students from different backgrounds as stepping stones for further learning;
- developing co-operative learning methods that create roles of high status and responsibility for each student in the class and that establish a positive peer climate for learning (Cohen, 1986; Slavin, 1990).
- providing curriculum material that enables individual students in heterogeneous classes to work on the same lessons at their appropriate levels (Epstein and Salinas, 1992).

Educational segregation

28. Educational systems differ in their degree of *segregation*, particularly the degree to which background factors determine the type of schooling a student receives. The meritocratic paradigm can be described by its desideratum that positions in education and later on in society are to be gained on the basis of 'merit', rather than on the basis of one's social origins. Educational opportunity (and presumably outcomes) should not be influenced by factors such as social class, residential location or gender. The determination, therefore, of the extent to which countries differ in the degree of segregation in education, and particularly how much schools in the various countries contribute to differences in student performance related to socio-economic status, is central to PISA's equity brief. Furthermore, it will be of value to examine which features of schools are potential causes of large (or conversely, small) achievement gaps between privileged and underprivileged students.

29. Lee, Bryk and Smith (1993) view the composition of the student body as an external influence on the school organisation which influences student outcomes. The authors distinguish three indicators of the composition of the student body —ethnic composition, social class composition, and gender composition — that may be seen as proxies for human and fiscal resources (taking an economic point of view) or as proxies for beliefs, values and expectations (taking a communitarian point of view). In both cases the empirical evidence that there are effects of these aggregate measures on student outcomes (over and above the effects of the student-level variables) is clear. This line of research suggests that it will be important to construct indicators of student body composition and explore whether these student body composition variables have a relationship with student outcomes.

Management and financing

30. Over the last several years many countries have seen shifts towards greater degrees of school autonomy in decision-making, and an increase in school based management (see for example Chapman, 1990, and Abu-Duhou, 1999). It would seem important, therefore, to analyse *the extent to which educational systems are rather centralised or decentralised and to what extent this has an impact on student achievement*. Chubb and Moe (1990) propose a model of autonomous schools where decision-making occurs on the local level and a funding system that depends on the success in free competition as an alternative to bureaucratic controls in centralised public school system. However, evidence of a greater effectiveness of autonomous schools is not very strong (see Creemers, Scheerens, and Reynolds, 2000, p.291).

31. A related issue is the distinction between *public and private sectors* education. The public/private distinction was central to the work of Coleman, Hoffer and Killgore (1981) who found considerable differences in student performance between private Catholic and public schools in the United States. Their findings on achievement differences between public and private schools led to the development of Coleman's social capital theory and a rethinking of market mechanisms in education (Chubb and Moe, 1990). More recent research, however, has shown that if the social intake and previous achievement of students is controlled for then there

are only small differences between the outcomes of public and private schooling (see Willms, 1987).

32. In a number of OECD countries there is an on-going debate about the relative effectiveness of private and public education systems. PISA 2003 will provide some opportunity to examine how much public, private independent and private government dependent schools differ in their students' outcomes, and to what degree these differences can be traced back to differences in student intake and selectivity. Note however that PISA will not have access to prior achievement indicators for the students, so there will be limits to the level of adjustment for intake characteristics that can be attained.

33. According to Murphy (1990) monitoring behaviour as practised in instructional leadership includes testing programmes, feeding back data on student performance to staff, departments and school management, as well as the evaluation of school's success in achieving its goals. Effective monitoring of student progress and programs to improve school effectiveness may also play a role in explaining school differences in student performance (van der Werf, 1995; Reynolds and Teddlie 2000). However, there is also evidence of negative effects of very frequent testing on student performance (Mortimore *et. al.*, 1988).

School size and location

34. For administrative and budgetary reasons, secondary schools often tend to be large (more than 900 or 1000 students). Large schools allow for economy of scale and more efficient use of both material and human resources. They can invest more easily than smaller schools in costly equipment (such as gymnasiums, swimming pools, libraries, science laboratories, and ICT resources) and in specialised staff (such as qualified administrators and support personnel for special needs students).

35. However, there is consistent evidence from the literature (Meier, 1995; Lee and Smith, 1996) that smaller schools (from 300 to 600 students) tend to be more effective in terms of student achievement and attitudes towards school, particularly for disadvantaged students. In small schools, teachers find it easier to collaborate and develop common goals; teacher/student relations are often more positive; students develop a greater sense of belonging; discipline is easier to monitor; dropout rates tend to be lower; and potential difficulties are likely to be identified earlier and to be solved in less bureaucratic ways.

36. Large schools may therefore benefit from restructuring policies aimed at creating several semi-autonomous units housed in existing large buildings, where students will experience the same kind of direct and positive contact with their peers and school staff as in small schools. A variety of organisational models (referred to in the literature as '*School-within-school*' policies) have been have been implemented in a number of urban school districts in the USA. Their effects were often reported to be positive (Raywid, 1996).

37. *Location of schools* is a school characteristic related to school size, which is also often correlated with the socio-economic background of students: Inner metropolitan areas usually have higher concentrations of students with disadvantaged social backgrounds and there is a tendency for rural areas to be underdeveloped and poorer. Rural schools also tend to be characterised by scarce resources, but have more homogenous student bodies are more likely to be cohesive (see Teddlie, 1994a).

38. The relationship between urbanism and school performance is not a straightforward one: Williams *et al.* (1993) report for Australia that there is no rural disadvantage in education once family and community attributes are taken into account. However, there is evidence that rural poverty may have greater impact on educational outcomes than urban poverty (Phelps and Prock, 1991). Furthermore, it has been observed that suburban schools have more favourable learning conditions than inner urban schools (Purkey and Rutter, 1987; Witte and Walsh, 1990).

School resources

39. The economists 'Educational Production Function' (EPF) based model, which assumes that inequities in measurable and financially related inputs (for example teacher/student ratio, teacher salaries, number of volumes in library) are related to differences in student achievement, has been used extensively in the research on school effects. Search for the EPF has led to serious debates about whether or not school resources matter to students' achievement levels (e.g. Hanushek, 1994; Hedges, Laine and Greenwald, 1994a and b; Monk, 1992). Most of the evidence available from both strands of research is reviewed in Scheerens and Bosker (1997).

40. Postlethwaite and Ross (1992) used data from the IEA Reading Literacy Study to identify characteristics (e.g. school size, type, staffing, and school management) that influence school effectiveness in reading. In an exploratory study of these data, Siniscalco and Ross (1997) tested a *International Reading Resources Scale* using the following types of indicators to describe the schools' human and material resources:

- *Human Resources*: Student-teacher ratio, total instruction time, school principals' ratings about the availability of specialised staff, existence of teacher evaluation, extra class lessons in reading, special remedial courses, etc.
- *Material Resources*: school principals' ratings about the availability of sufficient classroom material, existence of school library, number of books in the library, etc.

41. Siniscalco and Ross (1997) developed a scale that proved to be reasonably consistent and stable across countries. Using this scale schools could be classified into four described levels of resources. Though this scale was able to explain a large part of the between-school variance, after controlling for the effect of average school SES, no independent effect of resources on achievement was observed .

42. The lack of success of the economist's model, with its focus on quantifiable resource variables, has been attributed to its under-specification, in particular through the omission of important process variables such as school climate and classroom behaviour (see Teddlie and Reynolds, 2000, p. 302). Nevertheless, school resources as specified in the EPF model are central to public policy debates and are malleable variables. In particular, issues surrounding the adequacy of funding to schools and the implications of that funding for school quality and effectiveness are central to popular educational debate.

Instructional context within schools

Learning time

43. Learning time is one of the most precious instructional resources and there is a large body of research evidence that shows that for any given subject matter, the amount of actual exposure to instruction differs significantly from school to school. Efficient use of time may provide the students attending certain schools with three or more times as much mathematics instruction, for instance, than those in poorly managed schools (Anderson, 1980; Hossler, Stage and Gallagher, 1988).

44. In many cases, the amount of time *allocated* to the various subjects depends upon decisions that are not taken at the school level. However, schools can often expand to some extent the time allocation by providing extra courses for specific groups of students (such as remedial or enrichment courses) or by encouraging regular assignment of homework, or by proposing optional activities (e.g. science or theatre clubs).

45. Most of all, school policies can significantly reduce the amount of instructional time lost, or spent in non-academic activities, by implementing strategies aimed at reducing time lost

through students' and teachers' absenteeism, dead time, procedural or administrative routines, disruptions due to indiscipline, and so on.

46. Increased allocation of instructional time usually shows low, but positive correlations with achievement. It appears to be more beneficial for low-ability students or for students with high levels of anxiety than for other students (Gettinger, 1984; Guida, Ludlow and Wilson, 1985). It is also more beneficial for highly structured subjects, such as mathematics or Foreign Languages, than for less structured ones, (Block, Efthim and Burns, 1989). Increases in the proportion of instructional time actively spent by students in learning activities ('time-on-task') show much higher correlations with achievement gains. Unfortunately, collecting measures of time-on-task would require direct observational methods that would not be applicable in PISA. However, indirect indicators can be used to capture at least some related information, such as time lost at the beginning of each course, or disruption time.

Student support policies

47. Research syntheses on students at risk and of schools with disadvantaged intakes (for example Rossi and Stringfield, 1995; Laguarda 1995; Brown, 1999; Borman *et al*., 2001) consistently report positive effects on students' academic engagement and performance of a number of support strategies. These include, in particular:

- Organising individual tutoring of at risk students by peers or by volunteer adults;
- Organising one-to-one or small group mentoring by experienced teachers;
- Establishing homework support programs;
- Improving school-home connection for low-income parents;
- Developing comprehensive links with community services in order to co-ordinate the interventions of a variety of health, social, educational, cultural and welfare agencies.

School and classroom climate

48. A major criticism of early school effects research was the exclusion of school and classroom processes: Brookover (1978, 1979) addressed this criticism by including measures of school climate as perceived by students, teacher and/or school principals. Though school climate variables were strongly correlated with school level outcomes, there was considerable multi-collinearity among these measures and family background variables. According to Scheerens and Bosker (1997, pp. 112ff.) the concept of school climate can be viewed as a synonym of school culture that involves variables such as student engagement, absenteeism, student conduct and behaviour, staff motivation, and the relationship between students, teachers and school.

49. Related to the overall school climate is the learning climate at the classroom level which in early research on school effects had been ignored (Teddlie 1994b). Classroom observation studies provided evidence of how general school environment factors influence instructional quality in the classroom as well as the effects of teacher behaviour and classroom practice on the overall performance of schools (Schaffer, Nesselrodt and Stringfield, 1994). Classroom climate indicators may involve relationships within the classroom, order, work attitude and satisfaction (see Scheerens and Bosker, 1997, pp. 123 ff., see also the section on 'Learning and Teaching strategies').

50. It should be noted that measures of school and classroom climate are not only predictors of student performance but that they can also be regarded as important outcomes which may be affected by a variety of school factors.

Summary of potential policy issues

51. Issues related to this theme include:

Issues related to aggregation phenomena

- In the various PISA countries, what is the proportion of total variance in student outcomes (achievement scores, self concept, engagement and attitudes towards school) that is between-schools variance, as compared to within schools variance? Are the patterns observed in PISA 2000 as regards this decomposition of variance confirmed by PISA 2003 results?
- How do PISA countries compare in terms of the between-schools and within-schools differentiation of the study programmes offered ? (i.e. do all or most of the schools offer a same undifferentiated program ? or do they offer academic and vocational programmes within the same school ? or does the country have different types of schools, each of which offers only academic or only vocational programes? To what extent is ability grouping used for mathematics education?)
- To what extent are these school structures associated with lower or higher between-schools vs within-schools variance in the characteristics of the students' intake (SES, gender, language spoken at home, country of birth, etc.) ?
- Is there, as was the case in PISA 2000, some indication that school systems with highly differentiated structures often tend to show larger dispersion of student achievement scores and stronger impact on achievement of the student background characteristics? Are there other characteristics of schools that can be shown to increase or reduce the relationship between student background characteristics and student performance?
- To what extent do gender differences in student performance vary across schools? If so, are there any characteristics of schools that can be shown to moderate the relationship between gender and student performance?

Issues related to quality of school climate and of educational environment

- How do the PISA countries differ in terms of the average quality of the instructional context in their schools, as reported by students and principals? Do, for example, indices such as teacher involvement, teacher morale, student/teacher relations, student sense of belonging or school disciplinary climate show higher mean values in some school systems than in others?
- To what extent are these characteristics of the instructional context related to better achievement (after controlling for the school intake)?
- Is there any indication that schools with better climate and better instructional context tend to enrol higher SES students than other schools?
- Are there countries where the disparities in the characteristics of the instructional context are much larger across schools than in other systems?
- To what extent does the ratio of total time spent in school versus time used for instructional activities show between-schools variations? Are these variations associated with differences in school mean scores?
- Is the average amount of homework time (as reported by students) associated with particular patterns of school characteristics (e.g. type of study programmes offered in the school, student body composition, principal's perception of teachers' involvement or teachers' morale)?
- Is there any evidence that schools with disadvantaged intake tend to implement more systematic support policies than other schools (e.g. remedial courses, tutoring and monitoring practices, higher indices of teacher supportiveness)? Are there across-countries differences in this respect?
- In countries that use grade-repetition, what is the between-schools variance in percent of below grade students? Do schools with unusually high percentages of repeaters have higher indices of disengagement?
- Do students' engagement indicators (attendance, sense of belonging, interest in learning) tend to show higher values in smaller than in larger schools?

Issues related to school human and material resources

- To what extent do PISA countries differ in terms of their average indices of school Human and Material Resources? Are those variations entirely dependent on differences in GDP and in percent of national budget devoted to education? Are they associated with differences in mean country scores?
- To what extent do countries differ in the way their Human and Material resources are distributed across schools all over the country? Do some systems show high degree of uniformity, while others appear to have very unequal distribution of resources? Do some systems appear to be more efficient in equalizing Human resources, while large disparities remain in Material resources, or the reverse?
- In particular, to what extent do schools with high SES intake tend also to benefit from better Human and Material resources than schools enrolling disadvantaged populations? Is this inequity in distribution more important in some countries than in others?
- Is there any evidence that rural schools tend to have less Human and Material resources than other schools (after controlling for school size)?

Issues related to school autonomy, school management and funding

- How do PISA countries compare in terms of locus of educational decisions? Is the primary role in school management played by national authorities, or by regional authorities, or by the school council / the school staff?
- Are there countries with large differences in school autonomy (with part of the schools that strongly depend on central or regional authorities for most of the decisions, and part of them entitled to more local decisions)?
- Do schools with more autonomy differ from other schools in terms of their sources of funding?
- Is there any evidence that more autonomous schools tend to have better human and material resources than other schools (and/or better achievement scores, after controlling for intake)?
- Is there any evidence that between-schools variance in achievement is higher in decentralized systems?
- How do PISA school systems compare in terms of public/private management of the schools? (i.e. what percentage of the students, in each country, is enrolled in public schools, or in private schools with public funding, or in independent private schools?)
- Do public schools receive funds from non-government sources, and to what extent?
- Are there differences in human and material resources available in public versus private schools (when controlling for possible differences in size and rural/urban location)?
- How do public and private schools compare in terms of autonomy? (i.e., is there any evidence that many decisions are taken at the school level in private schools, while public schools are more dependent on central or regional levels of authority?)
- Are there differences in the type of populations served by various types of schools? (e.g. do the private schools tend to enrol more high SES students than public schools, or particularly able students, or, on the reverse, students who show underachievement problems?)
- Is there any evidence that coeducation is more developed in public schools than in private schools ? More generally, do private schools tend to enrol more homogeneous populations than public schools, or do they show the same degree of 'comprehensiveness' or 'segregation' as public schools in a same country?
- Do admission/transfer practices differ significantly between public and private schools ?
- How do public and private schools compare in terms of school environment offered to their students? (e.g. disciplinary climate, press for achievement, student-teacher relationship, teachers' involvement in helping students who have difficulties, quality of school life, etc.)
- How do public and private schools compare in terms of achievement (both mean scores and dispersion), when controlling for the characteristics of the intake?

- Are there similarities in the patterns of differences between public and private schools across all countries (or in specific groups of countries), or does the profile and 'role' of private education tend to be unique in each country?

Implementation

52. School-level information in the PISA study can be obtained through three different sources:

- Information on school type (relevant in tracked educational systems) and some other school characteristics are available from national sampling frames.
- Responses from student questionnaires can be aggregated to the school-level.
- School-level information is collected through the school questionnaire, which is completed by appropriate school personal. The data in the school questionnaire will be collected through a written survey, with phone interview follow-ups where necessary.

53. The following indicators for grouping practices of student are proposed:

- The existence of different programmes at system- and school-level and their respective orientation (academic, vocational, pre-vocational) and level of instruction (which is not always reflected in the orientation).
- School principal's reports on within-school and/or within-classes ability grouping.
- School principal's report on selection and transfer practices.

54. The following indicators for aggregation effects are proposed:

- Social composition of school (by taking the average SES of students).
- Percent of female students and existence of single-sex schools.
- Percent of students attending a grade below the expected grade.

55. Indicators for school management and funding are:

- Private/public school ownership.
- Sources of financing
- School autonomy in decision-making (as reported by the school principal regarding different aspects of school management).
- Implemented programs of monitoring school effectiveness and student progress (assessment practices, external evaluation of educational objectives).

56. Information about school resources can be obtained by measuring the following indicators:

- Information on inadequacies of human and material resources as judged by the school authorities.
- Indicators of instructional quality and equipment (e.g. teacher-student ratio, instructional time, teacher qualifications, number of computers).

57. To analyse the effects of school size and location systems it is necessary to collect information on:

- Enrolment of students.
- Urbanism of school location.

58. Learning time on the school-level can be measured on the following levels:

- School-level information on instructional time for different grades and programmes.
- Aggregated student reports on their time spent on learning.

59. Information on student support policies can be obtained from the following sources:

- Existence of special instruction designed to meet needs of low and high achieving students (remedial, enrichment classes).
- Student reports on study time spent in special courses (remedial, enrichment).

- Existence of special activities at school to promote student engagement in learning (maths clubs, competitions).

THEME 2: TEACHING AND LEARNING STRATEGIES

Elaboration

60. Theoretical and empirical research on teacher instructional practices, student learning strategies and their impact on student achievement is extensive. Selecting for PISA 2003 a limited set of appropriate indicators is not an easy task. Given the design of PISA, it is suggested to give priority to:

- Exploring dimensions that might reasonably be considered as being *pervasive* characteristics of either the instructional context or of students' learning strategies. By pervasive characteristics the characteristics are meant that often appear to be shared, to some extent, by different grade levels and across classes in a given school, and are more likely than others to have lasting effects on achievement. They are also more likely to show possible differences in the educational cultures of the various PISA school systems.
- Reflecting mathematics literacy as the major domain in PISA 2003 by focusing the data collection on the instructional context in mathematics classes and on the strategies used by students when working on mathematics tasks, while ensuring theoretical continuity with the major constructs used in PISA 2000 questionnaires and *self-regulated learning strategies* international option.

61. Many of the general components of current teaching/learning theories can be traced back to Carroll's (1963) model, where the student's learning outcomes are considered as a function of

- the time needed for the student to learn, which is determined by I.Q, ability to understand the instruction presented, prerequisite knowledge and the quality of the instruction received, divided by
- the *time the student spends in learning,* which depends on *opportunity to learn*, i.e. whether the topics to be learnt were actually taught or not, on the *amount of time allocated to them by the teacher*; and on the extent to which the student is willing to engage in learning tasks, i.e. *effort and perseverance*.

62. Derivations from this model have contributed to the operationalisation of concepts – such as quality of education, understanding processes, opportunity to learn, engagement in learning tasks – into a variety of sub-dimensions (for example Bloom, 1976; Slavin, 1994; Slavin, 1996), and in establishing links with other theories – constructivist theory, cognitive sciences, or models of social interactions in teaching/learning processes (for example Huitt, 1995).

Empirical findings on general factors associated with instructional effectiveness

63. The relationships with achievement of the dimensions identified by Carroll, and their later various operationalisations, have often been confirmed by empirical evidence from experimental studies, from studies on especially effective schools, from evaluation of school improvement programmes and from large scale longitudinal surveys, thus pointing to the global robustness of the Carroll model. The available syntheses of relevant literature (e.g. Purkey & Smith, 1983; Brophy & Good, 1986; Levine & Lezotte, 1990; Sammons, Hillman & Mortimore. 1995; Scheerens & Bosker, 1997) report positive effects on achievement of the following dimensions:

- *Consistency of teachers' practices and shared vision of the educational objectives pursued in the school*. One important aspect of instructional quality is the extent to which the school staff is able to convey to students a clear message on the learning goals that they are expected to attain. School effectiveness is therefore partly dependent on the principals- and

staff's success in building coherence and consistency concerning academic goals, innovations, and teaching practices (Levine and Lezotte, 1990). Given the traditional autonomy of teachers, consensus and cohesion cannot be taken for granted in schools. In an analysis of a number of indicators describing the self-reported instructional practices of 900 grade 9 teachers Grisay (1997) found that, for most of the dimensions explored, the differences between teachers allocated to the same class (as opposed to the between-class and between-school differences) represented more than 90% of the total variance. Interestingly, higher rates of inconsistency were found in 'innovative' rather than in 'traditional' schools, and for variables describing teaching styles rather than for variables describing assessment or disciplinary practices.

- *Adapting instruction for students with different knowledge prerequisites, different proficiency levels, or different learning styles*. The research results on the effectiveness of teaching differentiation or individualisation practices are quite mixed. However, positive findings are reported for specific methods of within-class grouping, for peer and adult tutoring, for interventions matching students' learning styles and, more generally, for the use of learning tasks with multiple formats, multiple levels of difficulty and/or multiple presentation styles.
- *Teacher attitudes and behaviours that create a motivating and achievement-oriented climate in the classroom.* The effective teacher profile emerging from the literature is that of a teacher who shows enthusiasm; who puts emphasis on academic work; who expresses high expectations on the potential for academic progress of each student, and feels responsible for helping them achieve the goals set; who provides all of them with frequent opportunities for success and positive feedback; who establishes fair classroom rules, mutual respect and attention paid to everyone's opinions, and a co-operative rather than competitive working climate (Peklaj & Vodopivec, 1999).
- *Opportunity to learn*, defined as consistency between the topics and processes taught and those assessed in the instrument used to evaluate student achievement;
- *Instructional time;* that is, time allocated, time actually spent on task and homework time all correlate positively with achievement. Not surprisingly, the most significant effects are observed for time on task. Indicators of discipline problems and of absenteeism, which can be considered as proxies for time wasted on non-instructional activities, were found to negatively affect achievement. (Brookover *et al.*, 1979, Rutter *et al.*, 1979).
- *Quality of the curriculum material and of the teaching processes,* in particular those that can foster achievement by facilitating understanding and/or eliciting appropriate learning processes, and/or developing efficient learning skills. One example is structured instruction, which consists of clear formulation and communication to students of teaching goals, accurate sequencing of the material to be learnt, ample practice time on varied and challenging learning tasks, frequent feedback. Other examples of effective strategies include use of advance organisers, high order questioning, inquiry approaches, and training in metacognitive processes or in creative problem-solving strategies.
- *Quality of assessment practices*. Frequent use of corrective feedback and of assessment practices focused on monitoring students' progress are consistently mentioned in the school effectiveness literature as having positive effects on achievement. The marking system used by the teacher may be important in this respect. Rheinberg (2001) distinguishes between teacher assessments based on social norms (i.e., the student is compared with other students in the class), criterion-oriented assessments (i.e. a comparison with given subject content standards is made) and individual assessments (i.e. student current performance is compared to previous performance). He suggests that the latter is more effective, in particular in terms of students' motivation. A question on the marks received by the students from their <Test Language>, mathematics and science teachers was included in the PISA 2000 Student Questionnaire. Interesting differences were observed across countries: in some school systems, up to 90% of the students said that their marks were above the pass/fail score (suggesting that the school reports are mainly used to certify student's mastery of the objectives set for them by the teacher), while in others significant proportions of students mentioned marks below the pass-fail threshold (probably indicating a more normative use of marks).

Suggestions from research on effective mathematics instruction

64. A long history of investigation in effectiveness of mathematics instruction has consistently established the value of *teaching for meaning* practices that create a mathematics classroom context where students can construct meaning by establishing connections between mathematical ideas, concepts and skills with other mathematical ideas, with concepts in other subjects and with real world referents (Grows & Cebulla, 2000).

65. Tanner and Jones (2000), among many others, categorise mathematical understanding as either *relational* or *instrumental*. Instrumental understanding, which is favoured by algorithmic teaching methods such as having the students learning rules and methods by heart, or working on lots of similar examples for one concept, often leads to a limited understanding of mathematics. Students with an instrumental understanding of mathematics might be able to recall rules and routines quite quickly, but they might not be capable of transferring concepts they have learned to other contexts and to real-life problem solving or application. Learning relational mathematics by 'problem-solving teaching' consists of building up a conceptual structure, or schema, from which the student can transfer solutions to different domains. Students are asked to explain their way of solving mathematical problems, and to learn how to transfer and apply new concepts. The authors provide evidence that, amongst other factors, particularly successful schools tend to emphasise relational rather than instrumental understanding.

66. Evidence from international surveys suggest that differences in the extent to which relational rather than instrumental/algorithmic teaching is used in mathematics classes should be considered as a possible explanation for both within-country and between country differences in mathematics achievement. Analysing the TIMSS-Repeat results, Mullis *et al.* (2000) showed that students in classes emphasising reasoning and problem-solving had higher achievement than those in classes with a low emphasis on these activities.

67. Based on the TIMSS videotapes collected in 231 grade 8 classrooms, Stigler *et al.* (1999) compared the dominant teaching styles used in Japan, Germany and the USA, and suggested that, while both in Germany and in the USA the lessons are structured in order "*to teach students a method for solving the example problem(s*)", in Japanese teachers' view "*understanding mathematics is the overarching goal; problem solving is merely the context in which understanding can best grow*". Thus, during a typical lesson in the US and Germany, the teacher instructs students in a concept, solves example problems with the class, then has students practice on their own, while he assists individual students. In contrast, the typical Japanese teacher starts by posing some thought-provoking problem; students work on it and present various ideas or solutions to the class. The class conclusions are summarised by the teacher, and similar problems are then practised by the students. Procedural routines represent about 90 % of the class time in the US, but only 40 % in Japan, where 45 % of the class time is spent in inventing new routines and analysing new situations.

Student's learning strategies

68. Teaching strategies and instructional practices are expected to affect achievement by promoting the development of effective learning processes among students, and their enduring interest and active engagement in learning tasks. Exploring to what extent possible national or international differences in some of the instructional dimensions mentioned above are associated with differences in students' learning attitudes and behaviours may therefore be of particular interest in PISA.

69. The area of research on learning strategies and self-regulated learning has been particularly productive over the last two decades or so, resulting in a number of different theoretical approaches (Zimmerman & Schunk, 1989, Baumert *et al.* 1998a). Usually, three main types of learning strategies are identified in the literature: strategies of information

processing (cognitive strategies), control strategies (meta-cognitive strategies), and resource-management strategies (Friedrich, 1995).

70. *Cognitive strategies* include *memorisation/rehearsal* (reading the material several times, learning important concepts, using mnemonic techniques), *elaboration* (constructing meaning, integrating, transferring) and *transformation* (transfer of information to another medium). *Meta-cognitive strategies* include *planning* (e.g. identifying learning goals), monitoring (checking whether the material has been understood) and regulation (adapting learning activities to the task). *Resource-management strategies* are related to identifying whether and when specific efforts should be provided to process a task.

71. PISA 2000 included an international option on *self-regulated learning as a cross curricular competence* (CCC), developed by a group of experts appointed by the OECD/INES Network A (Baumert *et al.,* 1998b). It included, in particular, two self-reported scales related to cognitive strategies (*memorisation/rehearsal* and *elaboration strategies*), one meta-cognitive scale (*control strategies*), and one scale related to action control (*effort/perseverance*), all adapted from existing instruments (the *Kieler Lernstrategien-Inventar* [KSI, Baumert, Heyn, & Köller, 1992], and the *motivated learning strategies questionnaire* [MLSQ, Pintrich *et al.*, 1993]). All scales, except memorisation, showed modest, but significant and internationally consistent correlation with students' achievement scores in all three domains.

72. However, both the KSI and the MLSQ instruments have a general orientation, thus the items included in PISA 2000 did not refer to learning strategies used in specific subjects, such as reading, mathematics and science. For PISA 2003, it has been suggested that there are conceptual and empirical reasons to believe that domain-specific indicators of self-regulated learning may be more closely related to academic achievement and may predict cumulative long-term achievement gains, in contrast to more global indicators (Baumert *et al.*, 2000). As a consequence, they recommended that the scales be maintained in the international option, but that all questions be reformulated, and focused on learning strategies when studying mathematics.

Summary of potential policy issues

73. Issues related to this theme include:

- How do the PISA countries differ in terms of the characteristics of the instructional context in their schools, as measured by selected indicators on time spent in learning, classroom climate, and teaching practices?
- Is there any evidence that certain types of teaching strategies have an impact on the students' learning processes or attitudes? For example:
  Do students taught in mathematics classes where problem-solving oriented teaching is practised show better elaboration and control strategies?
  Are students frequently exposed to co-operative group work more likely than others to express preference for co-operative learning situations?
  Are possible differences in teacher supportiveness or in assessment practices associated with differences in students' self-concept, interest in the subject taught, or engagement in learning tasks?
- Which of the characteristics of the instructional environment appear to be associated with significantly higher levels of achievement in all or in most of the PISA participating countries?
- Is there any evidence that certain types of instructional contexts or of teaching strategies are more effective with particular groups of students (e.g. disadvantaged students)?
- How large is the within-school variance of the characteristics of the classroom instructional context, as reported by the students? To what extent is lack of consensus among teachers (as reported by the principal) associated with larger variations in actual practices?

- How large is the between-schools variance in classroom instructional practices and in students learning strategies? Is there any evidence that schools that serve specific groups of students (e.g. high SES schools, or private schools) are more effective in implementing "good" teacher's practices and in developing "good" learning strategies among their students?
- How do the PISA countries compare as regards possible indicators of variation in the distribution of educational practices? Can any system be identified as particularly "homogeneous" in this respect, and if so, are there similarities between the institutional structures of systems with high vs low levels of instructional consistency across teachers and schools?

Implementation

74. Various sources can be used in PISA 2003 to collect information on the dimensions related to teaching and learning strategies. The Student questionnaire and the self-regulated learning questionnaire are the most appropriate instruments to collect self-reported information on students' learning strategies, attitudes and behaviours, as well as their perception of the instructional environment in their class. The School questionnaire can be used to cross-check students' answers on some characteristics of school context (e.g. school climate) and to collect information on principal's perception of consistency in school objectives and teachers' practices.

75. It is suggested that a new instrument: a *student information sheet* (to be completed by the school administrative staff), be added to provide information on the type of mathematics program attended by the sampled students and on the associated amount of maths instruction they receive.

76. Currently, the variables and constructs considered for inclusion in the PISA 2003 data collection are as follows:

- Teaching practices and instructional context
  Type of mathematics courses (e.g., general, advanced, remedial) offered in the student's study programme (included in the Educational Career questionnaire);
  Amount of instructional time in mathematics:
  Number of hours of instruction per school year in the various study programmes offered by the school;
  Time allocated to mathematics in the study programme attended by the student;
    Weekly time spent in mathematics courses;
    Homework time in mathematics;
    Time spent working with a mathematics tutor (if any).
  Assessment practices:
    School assessment policies;
    Mark in mathematics from the student's last school report (included in the Educational Career questionnaire);
  Item about the use of textbooks and other printed material in Mathematics classes;
  Mathematics classroom climate, as perceived by the student:
    Disciplinary climate;
    Perceived Teacher support;
  School climate, as perceived by the principal:
    School disciplinary climate (6 items, in School Q.);
    Teachers' morale (4 items, in School Q.);
  Consistency of school objectives, as perceived by the principal:
    Staff participation in setting school objectives;
    Staff consensus on school objectives;
- Learning strategies and behaviours used when working on mathematical tasks, as reported by the student (see a more detailed description under Theme 3):

Memorisation/Rehearsal strategies;
Elaboration strategies;
Control strategies;

THEME 3: STUDENTS' ENGAGEMENT IN MATHEMATICS

Elaboration

77. Students' engagement with learning is crucial for the acquisition of proficiency, and is also an important outcome of education. The relationship between engagement and achievement is almost certainly a reciprocal one: The more engaged students are in the process of learning, the more they will tend to learn, but levels of proficiency may also influence the level of engagement. Engagement of students may also be regarded as an aspect of general well being: Students who are engaged with their learning will tend to have less stress, and fewer social and psychological problems (Boekaerts, 1993).

78. There are a number of different forms of student engagement: *Engagement with school* reflects to what extent students have a positive relationship with the school and are involved in school activities. *Academic engagement* comprises students' general attitude toward learning and the extent to which they are engaged in learning activities. *Engagement in a subject* is the most specific form of engagement and reflects the extent to which students are dedicated to reading, mathematics, science or other subjects at school. In this theme *Engagement in mathematics* will be discussed. More general forms of student engagement, which are important components of school and classroom climate, will be dealt with in the theme *School Characteristics, Organisation and Structure*.

79. Engagement in a subject comprises a variety of factors that are often closely related to each other (*see* Guthrie and Wigfield 1997). It does not only refer to students active involvement in learning, but also to students' beliefs about their own ability to succeed in a subject, their motivation to learn a subject, their emotional relationship with a subject as well as their choice of learning strategies for a subject. This theme, therefore, will address the following basic questions regarding engagement with mathematics:

- *Self-related cognition*: Do students believe they can succeed in mathematics?
- *Motivational preferences*: Do students want to learn mathematics and why?
- *Emotional factors*: How do students feel about learning mathematics?
- *Behaviour-related Variables*: How do students learn mathematics?

Self-related cognition

80. Within the broader concept of self-regulated learning, self-related cognitions in general are supposed to have a considerable impact on goal setting, strategy use and achievement (Zimmermann 1999). Generally, self-related belief systems can be grouped into types: beliefs of *self-efficacy*, *control beliefs*, and *self-concepts*.

81. *Self-efficacy* - that is, student's "judgements of their capabilities to organise and execute courses of action required to attain designated types of performances" (Bandura, 1986, p. 391) - is deemed to have a strong influence on individual choices, efforts, perseverance and emotions related to the tasks. Self-efficacy can be regarded as one part of a comprehensive personal theory about the learner's own learning process, which directs his or her own learning (Bandura 1993).

82. In educational research, self-efficacy has often been measured through a very general self-assessment of an individual's capacities to master a subject. Such self-assessments require students to respond to questions that tap their general beliefs about being able to master a subject. This approach might be seen as one that assesses the students' *confidence* in having the necessary resources to succeed in a particular domain (see Baumert *et al.* 2000, p. 15).

However, the problem with these measures is that they are often generated *'without any clear academic activity or task in mind'* (Pajares 1996). Students who consider themselves proficient in mathematics in general, may give different answers when confronted with more specific tasks.[2]

83. Bandura (1986) stated that self-efficacy plays an important role in determining behaviour - that is feelings of confidence about a specific problem are crucial to an individual's capacity to solve that problem. Research has generally confirmed a relationship between mathematics self-efficacy and student performance, though different sizes of correlation were reported, often depending on the types of self-efficacy measures that were used (Multon, Brown and Lent, 1991). Hackett and Betz (1989) found that (task-specific) mathematics self-efficacy was a better predictor of career choice than test performance.

84. Research on *control beliefs* focuses on expectations by students of their ability to successfully master tasks and problems. Skinner, Chapman and Baltes (1988) distinguish between *strategy beliefs*, i.e. expectations about what is necessary to be successful, *capacity beliefs*, i.e. expectations about the one's own general ability to succeed (self-efficacy), and *control beliefs*, i.e. expectations about whether or not oneself is able to succeed, without reference to any specific means or capacities. Maximisation of all three processes is seen as relevant for producing optimal levels of engagement in learning (see Skinner, Wellborn and Connell, 1990).

85. Positive *Self-concept* can be seen as a desirable outcome variable of education (Branden, 1994), and its enhancement is one of the goals of policy-makers. Marsh and Shavelson (1985) argued that self-concept is multi-dimensional construct and distinctions between verbal, mathematical and (general) academic self-concept are required. The so-called *Internal/external frame of reference* (I/E) model (see Marsh, Byrne and Shavelson, 1988) assumes that students evaluate their own performance through social comparison processes; that is, their evaluation is based on their position relative to other students (external reference) and their relative performance on different school subjects (internal reference). Furthermore, it has been observed that school average of achievement tends to have a negative effect on self-concept; that is, students with same proficiency levels often have different levels of self-concept depending on the overall performance of a school (Marsh, 1990).

86. Recent research using the PISA 2000 data has shown (Marsh and Hau 2002), that in most countries individuals consider themselves either as verbal- or mathematical-type learners, a self-evaluation that does not always correspond to ability levels. Whereas there is a strong positive correlation between reading and mathematics proficiency, the correlation between verbal and mathematical self-concept is considerably lower or even negative in some countries whereas the correlations between mathematics proficiency and verbal self-concept and between reading proficiency and mathematics self-concept are generally negative.

Motivational preferences

87. Motivating students to learn mathematics is crucial in the process of fostering their engagement. United States national assessment data have evidenced that whereas primary students still tend to enjoy mathematics, attitudes toward this subject change dramatically among students in secondary education (Dossey *et. al.*, 1988).

88. Motivational preferences can be viewed as either 'intrinsic' or 'extrinsic' — students can be motivated either through internal (e.g. interest) or external reasons (e.g. perception of importance). Whereas intrinsic factors are related to the individual's interest or enjoyment of a subject, extrinsic factors are incentives for learning such as anticipated future rewards (for an overview see Middleton and Spanias, 1999).

[2] As an example of a task-specific efficacy scale see Kranzler and Pajares, 1997

89. An intrinsic motivational preference is *subject-related interest,* which affects continuity and intensity of engagement with learning, independently of the general motivation to learn (see an overview on interest research in Baumert and Koeller, 1998). Closely related to the interest dimension are enjoyment of mathematics and value of mathematics (see Aiken 1974). Intrinsic motivation is viewed as having positive effects on time on task, more comprehensive learning strategies, performance and activity choices in the absence of extrinsic rewards (Lepper, 1988). There is evidence suggesting that intrinsic motivation to learn is (at least partially) influenced by teacher supportiveness and classroom environment (see Middleton and Spanias, 1999).

90. In longitudinal studies *instrumental motivation* has been found to be an important predictor for course selection, career choice and performance (Eccles, 1994; Eccles and Wigfield, 1995; Wigfield, Eccles, and Rodriguez, 1998). Both intrinsic and extrinsic motivational preferences determine how much time and energy students invest in learning as well as choice of courses and learning strategies (see Deci and Ryan 1985).

Emotional factors

91. Emotional factors play a critical role in the learning and teaching of mathematics, because students perceiving mathematics as too difficult in view of their own self-esteem in this domain may tend to avoid the subject when possible (McLeod, 1992). Mathematics anxiety is the construct that has received most attention in this area; it is concerned with feelings of helplessness and emotional stress when dealing with mathematics. Though some researchers have treated this construct as part of general attitudes toward mathematics (e.g. Aiken, 1960), the majority have viewed it as distinct from attitudinal variables. Mathematics Anxiety is usually found to be negatively associated with achievement but this relationship can change depending on the students' social and academic background (Ma, 1999). It could also be shown that mathematics Anxiety has rather indirect effects on achievement, once self-related cognitions such as as self-efficacy and self-concept are taken into account (Meece, Wigfield and Eccles, 1990).

92. Another emotional factor is *positive emotional experience ('flow'),* which reflects the extent to which students are emotionally attached to learning mathematics and experience learning this subject as an enjoyable and meaningful activity (see Nakamura, 1988). This construct is closely related to subject-interest in mathematics (*see* Baumert 2000, p. 14). There is evidence that mathematics instruction that emphasises understanding of mathematical concepts leads to less anxiety and higher levels of enjoyment than teaching that stresses rote learning and is perceived as authoritarian (Middleton and Spanias, 1999).

Behaviour-related variables

93. Students may develop different types of learning strategies that shape their learning behaviour. Some main cognitive strategies are *memorisation* (learning key terms, repeated learning of material etc) and *elaboration* (making connections to related areas, thinking about alternative solutions etc.). C*ontrol strategies* are meta-cognitive strategies that involve planning, monitoring and regulation.[3]

94. Learning behaviour is also influenced by the students' preference for learning situations: Here, preference for co-operative learning for example, learning in groups, (Marsh, 1999) and preference for competitive learning, for example striving to be better than others (Owens and Barnes, 1992) are the most salient aspects. Cognitive and non-cognitive benefits of co-operative goal structures have been investigated in the past. Slavin (1983) showed in a meta-analysis of studies in this field that (task-specific) co-operative learning methods per se do not

[3] The strategies are collapsed because of a very high inter-correlations between them (Baumert *et.al.* 2000, p. 18)

affect achievement. However, co-operative learning including both individual accountability and group rewards/goals were reported to have positive effects on achievement. In PISA 2000 students that preferred co-operative learning methods tended to perform better than other students (OECD, 2000, p. 114).

95. Volitional processes also may have an impact on learning behaviour, i.e. students will have different levels of action control in addition to the motivational factors influencing their actual behaviour. One example of a volitional factor is the level of effort and perseverance students invest in their learning. Effort and perseverance are often difficult to distinguish and are therefore often regarded as two aspects of the same latent factor. Within the framework of self-regulated learning, this factor is regarded as one important component behaviour (Baumert *et al.* 2000; O'Neil and Herl, 1998). Schunk (1990) suggests that the persistence of students on difficult tasks is due to the belief that effort enhances their ability, i.e. their 'strategy beliefs' about what can be achieved through which means (*see also* Skinner, Wellborn and Connell, 1990).

96. Active involvement in the learning process has different connotations: Clearly, the *time spent on learning* is one crucial factor. Research has shown that time allocation (i.e. quantity of schooling) itself does not show a strong relationship with achievement (Anderson, 1980, 1983, 1985; Blai, 1986; Fisher and Berliner, 1985; Karweit 1976; Frederick and Walberg 1980). However, there is evidence that there is a moderate positive association between measures of specific 'time-on-task' (e.g. learning time spent on a subject) and student achievement (Guida, Ludlow and Wilson, 1985).

97. Large differences were observed with PISA 2000 data across countries in the overall amount of time devoted to *homework* in mathematics and other academic subjects (OECD, 2001). However, (i) there was no clear across-country relation between the average achievement in mathematics and average amount of homework, and (ii) within countries, the relation was mostly non-linear: students reporting less than 1 hour of homework had lower scores than those reporting 1 to 2 or 2 to 3 hours, but in almost all countries those reporting more than 3 hours of homework also had lower achievement scores than the intermediate categories. These inconsistent findings are probably also due to different homework policies depending on the educational systems: In some educational system larger amounts of mathematics homework may indicate instructional emphasis on this subject, or may indicate high levels of student interest and motivation, and correlate positively with performance. In other countries more homework may rather be an indication of remedial learning time for low-achievers.

98. Learning time is not the only component of active student involvement in the learning process, student participation in non-compulsory activities related to mathematics or choice of course combinations with an emphasis on this subject might be other *indicators of engagement*.

Concepts of engagement in mathematics

99. The various concepts that were discussed that might be relevant to the engagement of students in learning mathematics are related to each other. In fact, often it is an empirical question as to whether some of these concepts should be treated as separate or rather as single constructs (e.g. 'enjoyment of mathematics' and 'Interest in mathematics'). Previous research suggests that the relationship (a) between most these constructs and (b) of most of these constructs with achievement is reciprocal. Measures of self-esteem play an important role for developing an intrinsic motivation and have an effect on the development of emotional stress or positive emotional feelings. All of these constructs are also important for students' choice of learning strategies, time on task, choice of learning activities etc. and - in consequence - for academic performance. However, self-related cognitions, motivational preferences, emotional factors and learning behaviour may also be affected by achievement.

100. The concept of self-regulated learning, which includes many of the constructs described above, has received increased attention by policy-makers and educators. It postulates that self-related cognition, motivational orientations and learning strategies/behaviour are mutually dependent on each other (Boekaerts, 1997 and 1999, Baumert *et al.* 2000). The inclusion of these constructs in PISA 2003 will enable researchers to analyse their relationship in a comparative perspective and the correspondence of learning strategies/behaviour to motivational orientations and self-related cognition.

101. McLeod (1992) suggests that motivational research should focus more on the socio-cultural context, that is, the social organisation of school (for example 'sense of belonging'). This leads to the question to what extent subject-oriented engagement is related to students' general engagement at school and to what extent engagement in learning mathematics depends on the school's overall capacity to foster a positive learning climate (see 'School Characteristics and Student Performance').

102. Table 1 summarises the constructs within each of the broader domains.

Table 1: Domains and constructs related to students' engagement in mathematics

Domain	Construct
Self-related cognitions	Self-efficacy Subject confidence Control expectations Self-concept mathematics
Motivational factors	Subject-related interest Instrumental motivation
Emotional factors	Mathematics anxiety Positive emotional experience ('Flow')
Behavioural variables	Learning strategy: memorisation Learning strategy: elaboration Learning strategy: control strategies Preference for co-operative learning Preference for competitive learning Effort and perseverance (volitional) Time spent on learning (compulsory) Time spent on learning (voluntary)

Summary of potential policy issues

103. The following policy questions could be addressed by PISA 2003:

Self-related cognitions

- PISA 2003 will allow for the first time an internationally comparative analysis of the relationship between self-related cognitions and mathematics achievement. Inclusion of different kinds of self-related cognitions (e.g. global measures of self-concept as well as more task-specific measures of self-efficacy) will be helpful in clarifying this relationship.
- Self-related cognitions have often been found to contribute to choice of educational pathways and careers. This is especially important with respect to possible inequities (e.g. by gender): Career choices by students may be made rather on grounds of self-perceptions about skills than on abilities.
- To what extent are differences in self-related cognitions reinforced by differential instruction and teacher feedback (marks)?

Motivational preferences

- Intrinsic motivation tends to be lower at the secondary level and students seem often to loose interest and enjoyment in this subject after primary education. To what extent do educational systems vary in this respect and to what extent do these differences depend on heterogeneous versus. homogeneous modes of delivery of mathematics instruction?
- Mathematics is clearly an important subject for the future educational and professional career options of students. To what extent do extrinsic motivational preferences differ depending on programme orientation and what is the effect of motivation and motivational preferences on achievement?

Emotional factors

- Students' avoidance of mathematics due to emotional stress is reported to be widespread. Dislike of this subject is developed early and is often viewed as a result of experiences of failure. Are levels of mathematics Anxiety lower in educational systems with differential instruction (tracking or streaming) in this subject?
- Classroom practice needs to be attuned to the needs and abilities of students). To what extent do positive or negative emotional experiences depend on teaching styles and teacher feedback (marks)?

Behavioural variables

- Over the past decades efforts have been spent on reforming mathematics education by placing more emphasis on solving non-routine problems and higher-order thinking as opposed to more traditional methods that focus on routine mathematical tasks. What is the relationship between the learning strategies employed and the learning outcomes? To what extent does comprehension-oriented learning (e.g. elaboration) versus rote-learning strategies (e.g. memorisation) and preferences for co-operative/competitive learning situations correspond to different kinds of mathematical instruction as received by the students? And are there differences between educational systems/countries and between different kinds of study programmes in countries with heterogeneous school systems?
- PISA 2000 has shown that the relationship between homework and achievement is rather complex and is influenced by characteristics of the educational systems. This underlines the need for a more comprehensive measurement of learning time components in the next cycle which should enable a description of the amount of time spent on different types of mathematics study cross-nationally and its relationship with achievement. Are there differences between educational systems regarding their ability to foster learning of mathematics? To what extent do study programmes and/or tracks differ with respect to the learning time on mathematics within countries? Are there differences between compulsory and voluntary learning time with respect to their relationship with mathematics literacy?

Implementation

104. Most of the questions and issues that arise under this theme were dealt with through the *self-regulated learning* international option in PISA 2000. However, two important changes from the approach taken in PISA 2000 are proposed. First, given the agreed centrality of students' habits, perceptions, and attitudes towards learning mathematics to the successful outcomes of schooling it is proposed that these issues take a more central role in the PISA 2003 context questionnaires and analysis plans.

105. Second, as noted by Baumert *et al.* (2000), it is expected that the more proximal to the domain of interest these kinds of measures are - that is, the more focused they are on specific mathematics-related behaviours - the more powerful they will be as correlates of processes and outcomes, and the more suitable they will be for a descriptive purposes as part of an international comparison.

106. Consequently, items measuring the most important constructs are included in the core questionnaire and are no longer an international option. To provide data that will allow us to explore the most important of these issues the PISA 2003 student questionnaire includes the following scales:

Self-related cognitions:

- Self-efficacy: An item-battery asking students to rate their confidence in doing eight different mathematical tasks. In PISA 2000 global measures of self-efficacy were included as an international option. Here, a more task-specific approach with a focus on mathematics as the major domain was taken.

- Self-concept: 5 items are used to assess self-perceptions regarding the individual's mathematical talent.

Motivational factors:

- Instrumental motivation, i.e. the student's perception about the usefulness of this subject for studies and career (4 items).
- Interest and 'enjoyment' of mathematics (4 items).

Emotional factors:

- Mathematics anxiety: A 5-item-scale shall measure feelings of distress with mathematics.

Behavioural variables:

- Learning strategies: Elaboration (5 items), Rehearsal (4 items) and Control Strategy (5 items).
- Preference for Learning Situations: Two scales on Competitive (5 items) and Cooperative Learning (5 items).
- Questions regarding students' time spent on learning.

A Questions about students' participation in other (non-compulsory) activities in mathematics.

107. Based on the experiences in prior studies, emphasis should be placed on the analysis of cross-country validation of constructs to assure that measures of attitudes, perceptions and beliefs are comparable across cultures and educational systems. The constructs will be validated through an analysis of item dimensionality and scaling properties across countries.

THEME 4: MATHEMATICS AND GENDER

Elaboration

108. Gender differences in achievement are ongoing equity related concerns in OECD countries, and as such, are central to PISA. Given the focus of PISA 2003 on mathematics it is proposed that particular attention be paid to the following gender-related issues:

- Gender effects on mathematics literacy.
- Student gender stereotyping in mathematics – differences in mathematics related attitudes, habits and self-perception.
- The impact of school structures and teaching practices on gender differences in student's mathematics achievement, and mathematics related attitudes, habits and self-perception.
- The impact of school structures and teaching practices on gender differences in students' career preferences.

109. An extensive literature is available on the relationship between gender and mathematics proficiency, including a substantial body of work based on data from international studies such as SIMS (Burton, 1990; Brusselmans-Dehairs & Henry, 1994; Ethington, 1990) and TIMSS (Fierros, 1999). In addition a number of less comprehensive cross-national surveys have examined gender effects on mathematics achievement and attitudes (e.g. Schildkamp-Kundiger, 1982; Ma, 1995) and a plethora of national (both cross-sectional and longitudinal) studies have been undertaken.

110. Syntheses of the gender differences literature (see, for example ETS Policy notes Oct. 1989, Fan 1997, Bosker 1994) indicate that the patterns in gender differences in mathematics are often a function of the mathematics sub-domain. For example, boys are often found to perform better than girls in geometry, in mathematical reasoning, in problem solving, in applying mathematics to everyday situations, while girls sometimes appeared to have better computational skills, or to perform better in questions involving mathematical knowledge. Results at the item level from PISA 2000 suggest that the *reading load* in the item may well

also be an important factor. Items that involve a minimal amount of reading tend to favour boys while those that require more reading favour girls.

111. Even in countries where the overall mathematics results do not show significant gender differences, it is often observed that the variance in the scores of boys is greater than that for girls. That is, more boys are often found in the group of students with very high scores (and more boys are also found in the group with very low scores).

112. While gender differences continue to be of concern, some commentators have noted a decline in gender differences in mathematics over time (Tate 1997).

113. In addition to gender differences in achievement outcomes, gender differences have also been observed in research on self-related cognition, motivation and learning strategies: In general the mathematics self-efficacy of females has been reported as lower than that of males. Females, do however, invest more effort in learning tasks and show a greater preference for co-operative learning situations (Zimmermann and Martinez-Pons, 1990; Seegers and Boekaerts, 1996). Research has also shown that levels of mathematics anxiety of girls are typically somewhat higher than the mathematics anxiety of boys (Ma, 1999).

114. Even in the context of a narrowing gender gap in mathematics achievement questions may be asked about the remaining gender differences and how they relate to personal beliefs about mathematics and career choices involving mathematics (*for a discussion see* Fennema, 1996). Even when their achievement in mathematics is similar to that of boys, girls often appear to suffer from gender stereotypes — mathematics is considered as a 'masculine' subject. They may also experience lower levels of self-confidence or more anxiety in mathematics activities. This may well influence their decisions about enrolment in school tracks or study programmes where mathematics is an important subject. Which, in turn, may shape their post-secondary education and career choices.

115. Fennema and Sherman (1976) constructed a set of mathematics Attitude Scales (MAS) that have been widely used to explore the gender effects on mathematics learning. A key component of their research has been the role that is played by gender stereotypes. They have concluded that, the more a female stereotypes mathematics as a male domain — that is, the more likely she is to agree with statements such as 'Females need more help in mathematics than Males' — the less the less aptitude she would have to study and learn mathematics (see Forgasz, Leder and Gardner, 1999). An extensive meta-analysis of mathematics education research undertaken by Hyde, Fennema and Lamon (1990) concluded that while gender differences in perceptions of mathematics as a male domain had declined during the 1980s, they were persistent and still evident in current work. The authors presumed that stereotyping of mathematics as a male domain may still be critical to females' willingness to achieve in mathematics.

116. Cultural beliefs about gender (which may well be factors that influence stereotyping) are argued to bias individuals' perceptions of their competence at performing various career-relevant tasks. To the extent that individuals then act on gender-differentiated perceptions when making career decisions, cultural beliefs about gender channel men and women in substantially different career directions. Correll (2001) analysed the impact of gender-differentiated perceptions on career choice using data from the U.S. National Educational Longitudinal Study of 1988 (NELS-88). Her results show that males assess their own mathematical competence to be higher than their (equally mathematically competent) female counterparts, so even if they obtain the same grades and test scores, males are more likely than females to perceive themselves as mathematically competent. Interestingly, this bias in self-perception is not found for verbal tasks. Furthermore, feedback through school marks was found to have a larger positive effect on the mathematical self-assessments of females than those of males. Self-assessments of task competence were found to influence career-relevant decisions, even when controlling for measures of ability.

Summary of potential policy issues

117. PISA 2003 could address the following questions:

- What are the gender differences in mathematics literacy in PISA 2003 and do such differences, if any, vary across countries?
- To what extent do the gender differences in mathematics literacy in PISA 2003 vary across sub-domains and across item type (or format)?
- What can PISA say about gender differences in the variance of achievement levels for male and female students? In particular are female students over-represented in students who perform poorly enough to be considered at risk?
- Will the difference between TIMSS and PISA and/or the trend towards a reduction in differences between male and female performances be confirmed in PISA 2003?
- To what extent does PISA expose differences in mathematics related attitudes, habits and self-perception and do these differences vary across countries?
- Is there any evidence that school structures and teaching practices have an impact on gender differences in student's mathematics achievement, and mathematics related attitudes, habits and self-perception?

Implementation

118. The dependent variables needed to analyse the effects of gender on achievement, attitudes, occupational expectations and learning are discussed in connection with other themes in this framework. Students' gender is included in the questionnaire.

THEME 5: STUDENT'S EDUCATIONAL CAREER

Elaboration

119. One of the challenges faced by educational systems is to ensure that, although learning takes place in *collective* settings (schools, classrooms), the *individual* needs of learners are served in an efficient way. Many of the structural differences between systems can be attributed to the variations in policies aimed at finding an appropriate balance between two somewhat competing goals:

- giving all students in an age group equal opportunities to learn;
- adequately addressing the wide differences in the students' background, aptitudes, interests and learning profiles.

120. Considerable differences are found across countries with regard to the institutional aspects that are likely to shape the students' educational pathways:

- the extent to which pre-school education is delivered;
- the definition of special needs students and the extent to which they are enrolled in mainstream, or separate schools or programmes;
- the availability and the type of remedial mechanisms (e.g. withdrawal classes, special tutoring) used to help 'at risk' students;
- the use of grade-repetition and grade skipping;
- the number of schooling years spent by students in undifferentiated, 'fundamental' education, and the level at which curriculum differentiation occurs;
- how the differentiation is made: students may attend different study programmes either offered in different schools or in different tracks in the same school (*streaming* or *tracking*); they may attend a same programme, but in different classes, with different options and/or different levels of requirement for certain courses (*setting*); they may attend either flexible or stable groups in a same class, depending on their proficiency (*ability grouping*);
- how the decisions about the students' school career are monitored (e.g. through school-level versus district-level versus national-level assessments)

- the flexibility of transitions in the students' school life. For example, from primary to secondary cycle, from a type of programme or from a track to another programme or track, when changing school.

121. In a general sense, school systems can probably be arranged along a continuum, with one extreme being represented by countries where mainstream education is strongly privileged: students are kept in common, undifferentiated settings as long as possible – often until the end of compulsory full-time schooling. In such systems classroom-based mechanisms or practices are used to address differences in ability, background and interests in heterogeneous groups of students. Typically, no grade retention is used, and most of the special needs students tend to be integrated in ordinary classes. In such countries, the avoidance of an early specialisation is considered as important to ensure flexibility in future educational and professional choices, and to favour social integration and cohesiveness. Besides qualification, school is deemed to have an essential function in ensuring that students from various social, cultural or ethnic backgrounds learn to understand each other's culture and values, and in providing/eliciting the readiness to cooperate in a common society.

122. The other extreme would be represented by highly tracked school systems, where students are grouped into different types of homogeneous classes or schools, and receive differentiated programmes from an early age – often at the end of elementary school. Institutional mechanisms, such as high-stakes exams, are typically used to monitor the allocation of students to the various streams or tracks. Such policies rely on the idea that homogeneous settings will better serve the personal aptitude and interests of students, and that curriculum differentiation is important to provide students with specific qualifications linked to the various fields of occupation/work. In this respect, differentiated systems are often deemed to give the students a more "transparent" and "realistic" idea of their future educational and professional opportunities than other systems, and to produce a smoother transition from school to work (Hamilton and Hurrelmann, 1993; Hamilton, 1990).

123. The effects of heterogeneous as opposed to homogeneous grouping based upon student's achievement and self-concept have been (and still are) hotly debated. Indications from the educational literature (Slavin, 1990a, Slavin 1990b; Kulik and Kulik, 1992; Gamoran, 1995; Loveless, 1998) point to mixed results. Certain types of *within-class* or *cross-grade* ability grouping appear to be beneficial to both high-ability and low-ability groups of students. Controlled programmes (where the curriculum taught is kept constant across homogeneous and heterogeneous classes) typically appear not to produce significantly different effects. In most 'natural' settings, however, high-ability classes or tracks often appear to benefit from a more demanding curriculum, learning activities more often based on problem solving processes rather than on stereotyped drills, more qualified teachers, larger amounts of instructional time, etc. (Oakes, 1985). Significant gains were observed in the high ability groups, while for students with medium or low ability the results were either the same or poorer than for comparable students enrolled in heterogeneous classes. This suggests that tracking or streaming practices may result in some increase of the overall between-students variance in achievement.

124. Sociological studies (e.g. Alexander *et al.*, 1978; Bourdieu and Passeron, 1977; Lee and Bryk, 1988, Lee, 1993) provide evidence of the important role played by track and study programme pathways in mediating the effects of cultural and socio-economic status of the family on students' educational outcomes, aspirations and orientation in career choices. Drawing on a series of adult surveys conducted in the Netherlands (a highly tracked system) Van de Werfhorst (2001) concludes that the study programme chosen not only affects the student's educational attainment (in terms of number of years of schooling and level of qualification obtained) but study programmes also have a qualitative impact on the type of educational resources acquired – cultural capital, economic knowledge and skills, social skills, technological skills, which in turn are associated with differences in terms of values, motivational preferences and returns from the labour market.

125. Disruptive events in the student's educational career, such as long periods of illness or family moving to another location, can have negative effects on student's affective and social adjustment, and detrimental impact on his/her cognitive progress. Repeating grades, changing study programmes, and sometimes moving from a school to another in the same location can probably be considered as both a cause and a consequence of low achievement and is often associated with poor integration of the student and with problems in coping with school requirements. In particular, the school career of students who drop out before completing secondary school is often characterised by all sorts of disruptions, either in primary or in secondary grades, or both (Roderick, 1993). The group of future drops-out typically shows high percentages of students who were retained in grades, who changed schools two or more times, or who had long periods of absence from school.

Summary of potential policy issues

126. Issues related to this theme may include:

- How can education systems be classified along the dimensions of: comprehensiveness versus segregation (tracking), and early specialisation in certain fields of study/smooth transition to work versus long-lasting general education?
- Are there differences between countries with varying degrees of comprehensiveness/ segregation as regards:
  the mean level and the variability of achievement (and/or percentages of students in the highest and lowest levels of mathematics, science and reading literacy);
  the students' educational aspirations (after controlling for PISA achievement levels), and, in particular, the aspirations of low-achievers;
  the predictability of educational aspirations by PISA proficiency levels or by social class respectively (are aspirations mainly driven by achievement or by social background?);
  PISA achievement variance within schools.
- To what extent are the components of between-schools and within-schools variance attributable to students' pathways (e.g. to course taking patterns or to programmes of study)? In countries with tracked systems, what is the variance in PISA achievement in streams that are expected to be homogeneous?
- Is there any evidence that students in tracked systems, compared to students in un-differentiated systems, develop a more positive perception of utility of school for future professional career, and/or more realistic expectations in terms of their final level of education or their future job (in relation with their current study program and their level of proficiency)?
- Can the differences listed above (if any is confirmed) be typically attributed to specific segments of the school system? That is, are there specific study programmes, tracks or types of schools that appear to be responsible for most of the differences, or where the trends are better visible?
- If there are different programmes of study: what are the mean PISA achievement scores of students who attend each of them? Is the overlap in achievement of different tracks corresponding to the position of the tracks in the common tracking variable (P/V, see above)?
- What is the percentage of disrupted school pathways that are dependent on system structure (for example grade repetition, relegation in a lower track)? How does achievement of students with "disrupted" pathways compare to that of other students? Are there differences across school systems in the extent to which such "disruptions" are linked to social background?

Implementation

127. Only limited information was available in PISA 2000 about the student's educational career (current ISCED level; whether the student had repeated any grade; expected occupation at age 30). To explore this theme in more detail it is proposed that the variables in PISA 2003

should cover more components of the past, current and anticipated future educational circumstances of the student.

128. In particular, it would be essential to obtain more information on the type of study program attended. Despite its merits for an international study, the ISCED classification has limited descriptive power in this respect. For example, in a number of tracked systems, both "basic" and highly selective tracks in ISCED level 2 are classified as "type 2A" programmes, while there are huge differences in achievement between these tracks.

129. It is therefore suggested the following further information be collected.

- From school administrative records, the study programme attended by each sampled student, and the type of mathematics courses included in it;
- The grade level of the student and their study programme orientation (academic, general, vocational);
- The number of courses (if any) the student is taking at an "advanced" level. This information may help describing possible differentiation through setting mechanisms in comprehensive systems.[4]

130. Programmes of study in tracked education systems not only differ by their orientation in content (for example, vocational or academic), but, most often, also by the level of general ability or achievement of their students. When information on this explicit or implicit hierarchy is available, tracks may probably be classified in a way that is approximately comparable across different educational systems, using a formula developed by Ramseier (1997).

131. To the extent that tracking is dependent on general academic achievement, the extreme expectation would be that all students in a higher track show better achievement than any student in a lower track. Each track could therefore be described by its mean expected achievement (P) and by the expected variability (V) of achievement (heterogeneity) within this track. Assuming that each achievement level between a minimum and a maximum is equally probable in the population, the definition of mean expected achievement M=P of a specific track would be based on the percentages of students attending lower tracks in the system's population (p_{lo}) and the percentage of students in this specific track (p_{in}):[5]

$$P = p_{lo} + \tfrac{1}{2}\, p_{in}$$

Example: A system has three tracks: basic, general, academic, attended respectively by 20%, 50% and 30% of the students. Then:

P(basic) = 0 + ½*20 = 10; V= 20
P(general) = 20 + ½*50 = 45; V = 50
P(academic) = 70 + ½*30 = 85; V = 30

A completely heterogeneous system would have the values 50 / 100.

132. This type of variable was used in TIMSS and PISA 2000 national analyses, to compare tracks and types of schools across the 26 cantonal systems in Switzerland (where the tracking structures differ considerably). It explained large proportions of achievement variance in grade-based analyses. Developing similar indicators in PISA 2003 would require information from national statistics (or from the sampling frame) on the proportions of students attending the various relevant tracks, and information from the NPMs on which tracks are generally considered as "high", "medium" or "low" in terms of the expected ability of students enrolled in each of them.

[4] Number of courses at advanced level is used in Switzerland as a surrogate for tracking information when comparing comprehensive with tracked systems in different cantons.

[5] The variable *Mean expected achievement* could be improved by assuming a normal distribution of achievement and calculating the corresponding percentiles – but perhaps this would be an over-sophistication, given the crude assumptions on which the formula is based.

133. Additional information should also be collected on possible disruptions that occurred in the students' past school career, and on their educational expectations.

134. The following variables are therefore suggested for inclusion in PISA 2003:

- Past school career:
  Attendance of <ISCED 0> education;
  Age of entry in current school; and
- Possible disruptions that occurred in past school career:
  Grade repetition;
  Changing school (included in Educational Career questionnaire);
  Changing track (included in Educational Career questionnaire);
- Current institutional location of the student:
  Current grade, ISCED level and study programme orientation attended;
  Specific study programme attended (also collect through Student Tracking forms)
  Type of mathematics courses currently attended (basic, general, advanced) (included in Educational Career questionnaire);
  Mark received from the mathematics teacher in the student's last school report (included in Educational Career questionnaire).
- Expected educational level.
- Aspirations regarding future occupational status (included in Educational Career questionnaire).

THEME 6: USE OF AND ACCESS TO TECHNOLOGY

Elaboration

135. Competence with technology, especially computers, continues to grow in importance in terms of preparing students for success in their future lives. PISA provides an opportunity to explore both the level and nature of access that students have to Information and Communication technology (ICT), and the role that ICT plays in instruction. While more and more schools incorporate computer and other technology into the learning environment, there is an ongoing need for clarifying the role ICT plays in shaping the instructional context and how much it contributes to student achievement. The question of how new technologies (and here use and access to computer technology is perhaps the most salient issue) are used by students and how they are integrated into the educational system is of great relevance and is a recommended core theme for PISA.

136. For the purposes of PISA, ICT is defined as the use of any equipment or software for processing or transmitting digital information. The ICT with the greatest impact on people's life is the general-purpose computer and networks, especially the Internet. While these aspects of ICT will be a primary focus of PISA there will also be some investigation of the use of other technology, such as calculators, and its relationship with PISA mathematics proficiency.

137. It is proposed that use of and access to technology be analysed from three different perspectives:

- *The availability of ICT at schools*, that is, do students have access to ICT at school and to what extent does the school contribute to their knowledge and experience with ICT?
- *Students' ICT familiarity*, addressing whether students use ICT, what their level of experience and knowledge is and what their attitudes toward computers are?
- *The role of ICT in the instructional context*, that is, what role does ICT play in student learning and instruction?

138. Use of computer technology as part of education at school has become increasingly important and in many countries governments are promoting the integration of ICT into the

learning environment as a means for improvement. This position is driven by the assumption that new technologies are able to increase instructional quality and improve achievement.

139. Studies over the past 10-15 years have shown that the availability of computers at school has generally increased. The IEA SITES project, for example, has shown that at the beginning of the 21st century, computers are used in almost every school and that this trend goes beyond the major industrialised countries (Pelgrum and Anderson 1999). There does however, continue to be a concern that the expense associated with a high quality provision of an IT environment limits the access of such resources in those societies, and sections of society, that are less financially privileged — the so called *digital divide*. PISA provides an opportunity to undertake a detailed examination of the availability of ICT resources, the variation between countries and schools in that availability and the relationship between students' social background and their access to ICT.

140. Beyond simple availability of ICT the logical next step is to consider students' familiarity with ICT. Given the likely widespread availability of ICT (for the majority of students), PISA should take the opportunity to undertake an examination of familiarity that goes beyond frequency of use and access. While earlier studies have shown that large percentages of students had never used a computer (e.g. for American students Hicks, 1989), more recent studies report a substantial increase in students' computer use, especially in the United States and other industrialised countries (Pelgrum, Janssen, Reinen and Plomp 1993). PISA 2000 for example, showed that in most of the OECD countries[6] almost 100 percent of students have at least some experience in using computers, and an average of 65 percent of 15-year-old students stated that they were using a computer because they were very interested in it, and in no country was this figure less than 50 percent (OECD 2000).

141. Assessment of ICT familiarity (see an overview in Kirsch, Jamieson, Taylor and Eignor, 1998) should not now be limited to the use of computers - that is, variables such as frequency of use and years of experience (Geissler and Horridge, 1993; Powers and O'Neill, 1993). Rather the concept of ICT familiarity should be expanded to cover issues such as specific uses of ICT - that is, familiarity with certain software programs and procedures (Dalton 1994) or the places where individuals have access to ICT - school, home or other places. As Selwyn (1997) points out, one 'should be looking beyond finding out whether or not the student can use a computer toward how they are using the technology' also in particular, he recommends a focus on the creative use of ICT skills. A related area is the assessment of students' self-efficacy in ICT, i.e. their belief of being able to master tasks in this area (for an example see Fogarty et. al., 2001). Affective variables that focus on general attitudes toward this technology may also be important (Levin and Gordon, 1989).

142. While considerable resources are being devoted to the provision of ICT in schools, and there is a commonly held view that the use of ICT resources in instructional processes enhances the quality of instruction, evidence on the impact of technology use on student outcomes is sparse — particularly in the context of international studies . PISA provides an opportunity to take a macro perspective on this issue.

Summary of potential policy issues

- To what is extent does access to ICT resources vary between countries and between student and schools within countries. In particular, is there evidence that less financially privileged members of society have less access to ICT than do more financially privileged members of society?
- To what extent are there variations between countries and schools in levels of ICT familiarity, as measured through indices of frequency of use and self-efficacy?

[6] A total of 20 out of 32 countries used the ICT Familiarity option for PISA 2000.

- What are the relations between ICT familiarity, as measured through indices of frequency of use and self-efficacy, and student outcomes in reading, mathematics, science and problem solving literacy?
- Do some ways of using computers have a greater effect on achievement than others? What are the effects of calculator use on achievement, and how do these effects compare to the effects of computer use?
- What are the relations between ICT familiarity, as measured through indices of frequency of use and self-efficacy, and student background factors?
- Does the source of student's information and knowledge about ICT make any difference to self-efficacy, or to achievement?
- How much is this technology used at school in the PISA countries and what is the relationship between level of use and student outcomes?

B Is use of ICT associated with an increase in instructional quality? Is availability and use of ICT associated with particular aspects of the school instructional context?

Implementation

143. PISA 2003 will be limited to the use of written questionnaires for the collection of data on ICT. Information related to ICT will be captured from four sources. The most detailed data will be collected from the questionnaire to be used as part of the ICT international option. That instrument will be used to collect data on

- The availability of technology, and student access to it;
- The use by students of ICT, including both the kind of use and the frequency of use;
- Student attitudes to ICT, including their perceived self-efficacy with ICT; and

C The main source of students' knowledge about computers and the internet.

144. The school questionnaire will be used to elicit information about

- The total number of computers in the school, and

D the availability of computers and computer software for the purposes of mathematics and science instruction.

145. The student questionnaire will be used to generate information about:

E The existence of computers, educational software and internet facilities as part of students' home resources.

146. In addition, at the conclusion of every student cognitive test booklet, there is a question about student use of a calculator in answering the PISA cognitive test questions, and the type of calculator used.

147. In future PISA cycles it would also be desirable to use computer-based assessments of ICT experience and ability.

148. The school questionnaire for the Main Study will include questions about the number of computers at school and their availability to 15-year-old students.

149. For PISA 2000 an ICT familiarity instrument was used as an international option for participating countries. The PISA 2000 instrument consisted of 10 questions. Some of them were taken from an instrument developed by the *Educational Testing Service* (ETS) in response to concerns that the *Test of English as a Foreign Language* (TOEFL) may be producing biased results when administered on a computer, the bias resulting from different levels of computer familiarity. (See Eignor, Taylor, Kirsch and Jamieson, 1998). Some of the ETS items were slightly reworked for the PISA study and others were added to the PISA instrument.

150. Exploratory and confirmatory factor analysis showed that there were three clear dimensions: *Comfort and perceived ability with computers* (4 items), *Computer use* (6 items) and *Interest in computers* (4 items). The results show *that comfort and perceived ability with computers* was highest in the United States, Canada and Australia and lowest in Brazil whereas *Interest in computers* was very high also in countries with rather low levels of reported ICT ability among students. In view of the good reliability of these scales most of the items (some slightly modified) are included in the PISA 2003 ICT questionnaire.

151. In the PISA 2000 ICT questionnaire students were asked to rate their general ability to use computers. But global self-assessments only provide a very general self-perception that does not necessarily reflect an individual's experience and skills in this field. Therefore, similar to what is done in the case of mathematics Efficacy, an item battery was included where students are asked how confident they feel to do a range of different tasks with a computer. The tasks vary in their demand from very basic skills to more sophisticated computer knowledge and were included from a questionnaire used in the Australian Study 'Real Time, Computers, change and schooling' (Meredith *et. al.*, 1999).

152. Additionally, the PISA 2003 ICT instrument includes questions about how long students have been using computers, about the context in which students learned most about how to use computers (home, school, with peers) and about the sources of learning about the Internet.

THEME 7: FAMILY BACKGROUND AND STUDENT PERFORMANCE

Elaboration

153. Educational outcomes have been shown to be influenced by family background in many different and complex ways (Saha, 1997). For example, the socio-economic status of families has been consistently found to be an important variable in explaining variance in student achievement. However, differences in the use of measures have often led to quite different results regarding this relationship (White, 1982). In international studies additional caveats are imposed on the validity of background measures and there is a clear need for cross-national comparability of family background measures (see Buchmann, 2000).

154. In PISA 2003, the collection of information on family background – which was the focus component of PISA 2000 - will be reduced but it is proposed to maintain the following four aspects of family background:

- Socio-economic background, which includes occupational status, education and material possessions.
- Ethnicity, which includes immigration status and language use.
- Family structure.
- Family Social and Cultural Capital.

Socio-economic background

155. Socio-economic background is a particularly important variable for the PISA study for a number of reasons. First, the effect of socio-economic background on student achievement needs to be monitored over-time. It would be a major policy concern if its effects were increasing. It is important that identical measures can be constructed for each PISA cycle. Second, in many countries socio-economic background is a major reason for differences in achievement that might otherwise be attributed to family structure, or ethnicity. Third, the investigation of between school differences, school-level measures of SES can be derived from the students' SES.

156. Over the last 30 years a variety of measures of students' socio-economic background have been employed in research on educational and other outcomes. These measures are based on:

- parental occupational groups;
- parental occupational status or prestige;
- parental education;
- wealth and income; and,
- composites or combinations of the above.

157. Parental occupation is the basis for measures of social class although there are a number of alternative traditions of using class schema that have not met with convergence. In different countries, different sets of discrete categories have been, and are in use. Classifications that are in wide use tend to be amalgamations of Marxist conceptions, Weberian and Durkheimian ideas about the importance of ownership, education, and skill levels. One early consensus on cross-national comparability was reached by the work of Lipset & Zetterberg (1959) who saw that the large common denominator of national schemas was the division between non-manual, manual and farm occupations.

158. Parental occupational group has also been used widely in empirical research. A variety of measures of socio-economic background can be constructed from occupational data ranging from simple dichotomies distinguishing white-collar (or non-manual) backgrounds and blue collar (or manual) backgrounds, to more complex measures with 12 or more occupational groups. Such measures may or may not distinguish professionals from managers, clerical and sales workers from other groups, the trades from other types of manual work, and farmers from other groups. Furthermore, occupational data is used to construct continuous measures of occupational prestige and socio-economic status. Therefore occupational data is a very useful because of its flexibility.

159. Occupational Prestige is based on the idea that occupations can be ranked in terms of prestige or social status. Such measures are based on reasonably detailed occupational data. This prestige research tradition and its conclusions are solidly documented in 'Treiman's 'Occupational Prestige in Comparative Perspective' (Treiman, 1977). Treiman consolidated all existing prestige surveys and found the prestige of occupations is very similar across modern societies. Treiman averaged the hierarchies available to him in his Standard Index of Occupational Prestige Scores (SIOPS). He showed that SIOPS can predict criterion variables as well as local or national measures of occupational prestige.

160. One group of occupational status measures are the socio-economic indices (or SEI). These measures are based on the concept of occupational prestige. However, the prestige of an occupation was developed from the judgements of survey respondents. Since respondents cannot be asked about the prestige of all occupations, these studies could only rank the prestige of a limited number of well-known occupations. Therefore, there were perhaps hundreds of occupations without a prestige ranking. This problem was solved by Duncan who estimated the prestige of these occupations without a prestige ranking, by using a prediction formula, that was obtain by regressing the prestige of an occupation (for those occupations for which it was known) on the mean education and mean income of the jobs incumbents (taken from the US Census). Using the regression weights, prestige rankings could be assigned to all occupations. Duncan (1961) discovered that this worked well, not only for the occupations for which he tried to estimate a prestige score, but also for the occupations for which prestige was known. Duncan replaced the prestige measure by his new socio-economic index and other researchers adopted it too. Similar procedures were used by other researchers elsewhere and Duncan's procedure has been replicated in other countries and whenever national statistical agencies change their classification of occupations (Featherman & Stevens, 1982).

161. An internationally comparable SEI index (the ISEI = International Socioeconomic Index of occupational status) was created by Ganzeboom, de Graaf & Treiman (1992), who

proposed a new conceptualisation of SEI. Ganzeboom *et al.* conceive of occupational differentiation as the 'process that converts educational credential into earnings'. By implication, occupations can be scored as a scale that maximised the indirect effect of education on earnings and minimised the direct effect. Note that with this definition and procedure, there is no longer a link with prestige as a criterion variable and thus removes one of the ambiguities in the interpretation of SEI scores. Conceptual SEI can be understood as the resources required for socio-economic advancement in modern society. Therefore occupational status is a resource that people can use for the benefit of themselves and their children. Ganzeboom *et al* applied the procedure to data from 17 countries, at different levels of economic development and political order. As is the case with occupational prestige, there is little between country differences variations in the ranking of occupation and the international measure tends to be as effective as national measures on criterion variables. This SEI index was used in PISA 2000.

162. In the PISA 2000 differences in reading, mathematics and science literacy were found to be strongly associated with occupational status. In some countries like Belgium, Germany and Switzerland students in the bottom quarter of the occupational index were found to be more than twice as likely as other students to be in the bottom 25 per cent of their country's students on the reading literacy scale (OECD, 2001, p. 140).

163. Parental education is a second family background variable that is often used in the analysis of educational outcomes. Theoretically, it has been argued that parental education is a more relevant influence on student's outcomes than is parental occupation. Like occupation, the collection of internationally comparable data on parental education poses significant challenges, and less work has been done on internationally comparable measures of educational outcomes than has been done on occupational status. The core difficulties with parental education relate to international comparability (education systems differ widely between countries and over-time within countries) and response validity (students are often unable to reliably report their parent's level of education).

164. Early research suggested that mother's education has unique positive effect on achievement net of father's education and occupational status (Sewell and Hauser, 1975; Mare, 1981). In view of the high correlation between both measures for parents' education researchers normally use only mother's education or a combination of both.

165. As in the case of occupational status comparability of educational levels is a problem for international comparative research. Scales like ISCED or CASMIN have been developed to ensure cross-national comparability. PISA 2000 used the ISCED classification and found large differences between students whose mothers had completed upper secondary education and those whose mother had the lowest level of attainment (OECD, 2001, p. 149).

166. A third aspect of socio-economic status is wealth, a dimension tapped in international studies through ownership of material possessions. There is considerable evidence that income and wealth are related to student achievement and other learning outcomes, however, effects of material resources are generally considered to be weaker than those of parents' education (Feijgin, 1995; Ganzach, 2000). Additionally, Teachman's (1987) measure of educational resources in the home was found to have effects on test scores after controlling for other background factors.

167. PISA 2000 used an index of family material possessions as an indicator for wealth that was generally positively correlated with reading literacy. Notably, the correlation was highest in developing countries like Mexico or Brazil and in the United States, whereas in many other industrialised countries there was only a weak relationship (OECD, 2001, p. 143).

168. PISA 2000 also used a combined index of occupational status, parental education, cultural and material possessions, and home educational resources as a measure for the general socio-economic family background. So-called socio-economic gradients were employed to

describe the impact of socio-economic status on student performance cross-nationally (OECD, 2001, pp. 184 ff.).

Ethnicity

169. Ethnicity is defined here as family background characteristics such as country of birth (of student and parents), use of language or belonging to an ethnic minority in a country. International studies have confirmed differences in reading literacy for language and immigrant status (see Elley, 1992; OECD, 2001) but also for mathematics literacy (Mullis *et. al.* 2000). One rather obvious explanation for ethnic differences in student performance is that students from immigrant families, especially among those who have arrived recently, lack proficiency in the language of instruction. Alternatively, cultural factors can play a role here: unfamiliarity with cultural norms of the dominant culture may be an obstacle for learning at school. Furthermore, ethnic differences can also be explained with differences in socio-economic status, ethnic minorities often have a lower socio-economic status, which in turn is highly correlated with student performance. However, there is some evidence that immigrant status and language do have a unique impact on student literacy (Lehmann, 1996).

Family structure

170. Family structure represents an important factor of socialisation, which may affect learning outcomes. Research in the United States, for example, has shown that students from single-parent families perform less well than those from two-parent households. This typically has been associated with economic stress and lack of human or social capital in the household (*see* McLanahan and Sandefur, 1994; Seltzer, 1994). However, the effects of single-parent upbringing on learning outcomes have been generally considered as relatively small (see a review in Marjoribanks, 1997). The first report on the OECD PISA confirmed that reading literacy scores were significantly lower for students from single-parent families in a majority of countries (OECD, 2001, p. 152).

Family social and cultural capital

171. Families possess *social capital*, which consists in the relations among family members (Coleman 1988), and *cultural capital* or the knowledge of cultural values dominant in a society (Bourdieu and Passeron, 1977). Both concepts have become popular among educational researchers and are deemed to affect students' learning outcomes.

172. Coleman (1988) defined social capital as a social-structural asset that facilitates the success of an individual's actions and, thus, her learning outcomes. It may exist either within or between families, and relates also to parent-child relations such as the attention parents pay to children and their education. One example is the quality of *parent-child communication*, which is reported to correlate with student performance (see Howerton, Enger and Cobbs, 1993).

173. Families may vary in the degree of possessing cultural capital. Cultural capital is seen as those 'widely shared, high-status cultural signals (attitudes, preferences, formal knowledge, behaviours) used for social and cultural exclusion' (Lamont and Lareau, 1988, p. p. 156). Typically this concept is operationalised as parents' participation in *highbrow* cultural activities and possessions of cultural objects such as books, art, or music (Teachman, 1987). There is evidence that student or parent participation in, and preferences for, cultural activities as going to concerts, listening to music, or reading literature are positively correlated with student performance (DiMaggio, 1982; De Graaf, 1986). However, the concept has been criticised for its definition of highbrow culture, which may vary across country and makes it difficult to implement in a cross-national study (Farkas 1996; De Graaf, DeGraaf and Karykamp, 2000).

174. In PISA 2000 two scales on cultural and social parent-child communication were used, both scales showed a positive relationship with reading literacy, the correlation with

cultural communication being somewhat higher (OECD, 2001, p. 147). An index for the possession of cultural items was found to be also closely related to differences in performance, the correlation was generally higher than between the index of family wealth and student performance (p. 144).

Potential policy questions

175. PISA 2003 could address the following questions regarding the relationship of family background and student performance:

- To what extent is socio-economic family background correlated with student performance at school? What are the differences between countries regarding this relationship and to what extent are they influenced by the specifics of educational systems? Do these differences correlate with the other indices of social segregation at the country level?
- PISA provides a unique opportunity to monitor the relationship between family background and student performance over time: Are there any changes in the relationship between home background and student outcomes between PISA 2000 ad PISA 2003. If there are changes do they correspond to any changes in the educational systems?
- What are the consistencies and inconsistencies between PISA and other international studies, such as TIMSS) in the relations between home background factors and student outcomes?
- To what extent are students from immigrant families and ethnic minorities disadvantaged in terms of educational outcomes? Are there ethnic differences in participating countries and to what extent are these explained by language use and socio-economic background? Are there countries where these differences are smaller and are there countries where the integration of students from immigrant families at school is more effective? Here, trends analyses are also an important aspect of PISA, as policies that aim at a better integration of ethnic minorities into the educational systems can be evaluated with respect to their effectiveness.

Implementation

Socio-economic background

176. In PISA 2000, occupational data for both the student's father and student's mother was obtained. It was coded to a four digit ISCO codes (ILO, 1990) and then mapped to Ganzeboom *et al's* SEI index. It is proposed that this procedure be implemented again in PISA 2003. Consideration was given to reducing the response and data-coding burden by asking the occupation questions for father's occupational status only. It was noted, however, that in PISA 2000, both father's and mother's occupational status (measured by ISEI) had independent effects on student achievement. Further it was deemed important that consistency in the implementation of key variables such as ISEI be maintained over time. Occupational data are again collected for both parents in PISA 2003.

177. During PISA 2000 some questions were raised regarding the reliability of occupational data obtained from students. The validity of this data has been reviewed in a number of ways. First, the occupation distributions at the national level have been compared with external statistics, such as data from the European Labour Force Survey (EULFS) of 1999. These analyses have in general supported the validity of the PISA 2000 data. Secondly, data analyses using these data have shown credible results that are consistent with expectation. Third, a number of validity studies where undertaken with the PISA 2000 Field Trial. These validity studies (see Adams and Wu, 2002) support the accuracy of the occupational data. Interestingly one finding from these validity studies was that respondents and their parents often appear to be describing the same occupation but using slightly different descriptions, and these descriptions are then coded differently. It has been suggested that this problem might be alleviated somewhat by asking for more details from the student and by using two 'blind' coders, where differences in the coding are adjudicated by a third party.

178. Indices of parental education will be retained in PISA 2000, both for consistency with PISA 2000 and because it has been shown to influence educational attainment over and beyond the influence of parental occupation. It is again recommended that parental education be classified with ISCED (OECD 1999). Care needs to be taken in the collection of data however: A major issue with international comparability is the relative standing of parents with vocational education. Furthermore, combining different educational levels, which belong to the same ISCED category often obscures important differences: Sometimes, different educational tracks or programmes receive the same ISCED classification. Therefore, it is proposed to include additional variables for parental education that preserve the national educational levels in the international database in order to permit later secondary research on this issue. Furthermore, a different question format for school education is used in PISA 2003 to avoid the discriminatory character of the previous format.

179. Material possessions, at home, as an indicator for family wealth, was not a consistent predictor of student performance in PISA 2000, it is recommended that it be retained, and considerations be given to its improvement, however, to allow trends analysis.

Ethnicity

180. Both use of language and country of birth for student and parents needs to be retained for PISA 2003. One reason for including it is the consistency with PISA 2000, the other can be concluded from research showing that these are important correlates of educational performance. However, the question used in PISA 2003 is modified so that it can capture a richer set of data. In particular, countries were encouraged to ask for a wider range of countries of birth and languages. The responses will be coded using international country and language codes. Retaining data at this level will permit more in-depth study of the relationship between the integration of subgroups of immigrants and student performance.

Family structure

181. A simplified approach to family structure be taken in PISA 2003 compared to the one taken for PISA 2000. In particular, the question about other family members at home captures only details about parents/guardians because it provides sufficient detail to analyse differences in performance between students from single-parent and two-parent families.

Family social and cultural capital

182. The possibilities for the measuring of both social and cultural capital of students' families are limited in PISA: An adequate measurement of both concepts requires more information about student's parents than can be reasonably obtained through the student questionnaire.

183. The PISA 2003 Student Questionnaire includes an index of cultural possessions at home, and the Number of books at home.

REFERENCES

Abu-Duhou, I. (1999) School Based Management. Paris: UNESCO.

Adams, R.J and Wu, M.L. (2002) *PISA 2000 Technical Report*. OECD Publications. Paris

Aiken, L. R. (1974). Two scales of attitudes toward mathematics. *Journal for Research in mathematics Education*, 5, 67-71.

Alexander, K. L., Cook, M. A., and McDill, E. L. (1978). Curriculum tracking and educational stratification. *American Sociological Review*, 43, 47-66.

Anderson, L.M. (1980) Learning Time and Educational Effectiveness. *NASSP Curriculum Report 10*. (ED 210 780).

Anderson, L.M. (1983). Policy Implications of Research on School Time. *The School Administrator* 40, 25-28.

Anderson, L.M. (1985). What Are Students Doing When They Do All That Seatwork. In: C.W. Fisher and D.C. Berliner.*Perspectives On Instructional Time* New York: Longman.

Baker, D., Goesling, B., and Letendre, G. K. (2002). Socioeconomic Status, School Quality, and National Economic Development: A Cross-National Analysis of the "Heynemann-Loxley-Effect" on mathematics and Science Achievement. Comparative Education Review, Vol. 46, No. 3.

Bandura, A. (1993). Perceived Self-Efficacy in Cognitive Development and Functioning. *Educational Psychologist*, Vol. 28, No. 2, 117-148.

Baumert, J., Heyn. S. and Köller, O. (1994), *Das Kieler Lernstrategien-Inventar (KSI)*, Institut für die Pädagogik der Naturwissenschaften an der Universität Kiel, Kiel.

Baumert, J., Klieme, E., Neubrand, M.,Prenzel, M., Schiefele, U., Schneider, W. and Tilmann, K.J. (1998a). *Self Regulated Learning as a Cross Curricular Competence*. Berlin: Max Planck Institut for Bildungforschung.

Baumert, J., Fend, H., O'Neil, H.F. and Peschar J.L. (1998b) *Prepared for Life-Long Learning. Frame of Reference for the Measurement of Self-Regulated Learning as a Cross Curricular Competence (CCC) in the PISA Project.* Berlin: OECD, 1998.

Baumert, J., Klieme, E., Neubrand, M., Prenyel, M., Schiefele, U., Schneider, W., Tillmann, M., Weiss, M. (2000). Self-regulated Learning as a Cross-Curricular Competence. Berlin: Max-Planck-Institut fuer Bildungsforschung.

Baumert, J. & Köller, O. (1998). Interest research in secondary level I : An Overview.In L. Hoffmann, A. Krapp, K.A. Renninger & J. Baumert, (Ed.), Interest and learning. Kiel: IPN.

Beady, C., Slavin, R.E. & Fennessey, G. (1981). Alternative student evaluation structures and a focused schedule of instruction in an inner-city junior high school. *Journal of Educational Psychology,* 73, 518-523.

Blai, B., Jr. (1986). Educational Reform: It's About 'Time'. *The Clearing House* 60: 38-40.

Block, J.H.; Efthim, H.E.; and Burns, R.B. (1989) How Well Does Mastery Learning Work? In *Building Effective Mastery Learning Schools*. New York: Longman.

Bloom, B.S. (1976). *Human characteristics and School Learning.* McGraw Hill.

Boekaerts, M. (1993). Being Concerned With Well-Being and With Learning. *Educational Psychologist*. Vol. 28, 2, 149-168.

Boekaerts, M. (1997). Self-regulated learning: A new concept embraced by researchers, policy-makers, educators, teachers, and students. *Learning and Instruction*, 7(2), 161-186.

Boekaerts, M. (1999). Self-regulated learning: Where we are today. *International Journal of Educational Research*, 31, 445-475.

Borman, G.D.; Stringfield, S.C.; Slavin, R.E., Eds. (2001). *Title I: Compensatory Education at the crossroads.* Sociocultural, Political and Historical Studies in Education. Washington, D.C.: Office of Educational Research and Improvement (ED) (ED452322).

Bosker, R.J. (1994), School differences in producing gender-related subject choices, School Effectiveness and School Improvement, , v.5, 2, 178-195.

Bourdieu, P., and Passeron, J. C. (1977). *Reproduction of Education, Society and Culture.* Beverly Hills: Sage.

Branden, N. (1994). *Six pillars of self-esteem*. New York: Bantam.

Brookover, W., Beady, C., Flood, P., Schweitzer, J., & Wisenbaker, J. (1979). *Social school systems and student achievement: Schools can make a difference.* New York: Praeger.

Brophy, J., & Good, T., (1986). Teacher behaviour and student achievement. In M.C. Wittrock (ed.), *Handbook of research on teaching*. New York: MacMillan, 328-375.

Brown, D. (1999). Proven Strategies for Improving Learning & Achievement. Washington, D.C.: Office of Educational Research and Improvement (ED) (ED430179).

Brusselmans-Dehairs, C. & Henry, G. (1994), Gender and mathematics, International Journal of Educational, v.21, 4, 351-438.

Buchmann, C. (2000). Measuring Family Background in International Studies of Educational Achievement: Conceptual Issues and Methodological Challenges. Paper presented at a symposium convened by the Board on International Comparative Studies in Education of the National Academy of Sciences/National Research Council on November 1, 2000, in Washington, D.C.

Burton, L. (1990), Gender and mathematics : An International Perspective. London, Cassel.

Carroll, J.B. (1963). A model of school learning. Teachers College Record, 64, 723-733.

Chapman, J. D. (Ed) (1990). School based decision making and management. London: Falmer Press.

Chubb, J.E., & Moe, T.M. (1990). *Politics, markets, and America's schools.* Washington: Brookings Institution.

Cohen, E.G. (1986). *Designing groupwork: Strategies for the heterogeneous classroom*. New York: Teachers College Press.

Coleman, J.S. et.al. (1966). *Equality of educational opportunity.* Washington: D.C.U.S. Department of Health, Educationa and Welfare.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of sociology*, 94, S95-S120.

Coleman, J.S., Hoffer, T., & Killgore, S. (1982). *High school achievement. Public, Catholic and private schools compared.* New York: Basic Books.

Conti, S. D.; Ellsasser, C. W.; Griffin, G. A. (2000). *School Restructuring: A Literature Review.* Washington, DC: Office of Educational Research and Improvement (ED).

Creemers, B., Scheerens, J. and Reynolds, D. (2000). Theory Development in School Effectiveness Research. In: Teddlie, C. and Reynolds, D. (Eds.). *The International Handbook of School Effectiveness Research* (pp. 283-298). London/New York: Falmer Press.

Dalton, D. (1994). What do others know about CD-ROMs, LANs, Modems and More? A Survey for Staff Training, *The Book Report*, 12, 19.

Deci, E. L. and Ryan, R. M. (1985). Intrinsic motivation and self-determination in human behavior. New York: Plenum Press.

DiMaggio, P. (1982). Cultural capital and school success: the impact of status capital participation on the grades of U.S. high school students. *American Sociological Review* 60: 746-61.

Dossey, J. A., Mullis, I. V. S., Lindquist, M. M. and Chambers, D. L. (1988). *The mathematics report card. Are we measuring up? Tends and achievement based on the 1986 national assessment.* Princeton, NJ: Educational Testing Service.

De Graaf, N. D. (1986). The impact of financial and cultural resources on educational attainment in the Netherlands. *Sociological Education* 59: 237-46.

De Graaf, N. D., De Graaf, P. M. and Kraykamp, G. (2000). Parental cultural capital and educational attainment in the Netherlands: a refinement of the cultural perspective. *Sociological Education* 73: 92-111.

Douglas, J.W.B. (1964). *The Home and the School*. London: McGibbon and Kee.

Eccles, J. S. (1994). Understanding women's educational and occupational choice: Applying the Eccles et al. model of achievement-related choices. Psychology of Women Quarterly, 18, 585-609.

Eccles, J. S., & Wigfield, A. (1995). In the mind of the achiever: The structure of adolescents' academic achievement related beliefs and self-perceptions. Personality and Social Psychology Bulletin, 21, 215-225.

ETS (1989), The Gender Gap, ETS policy notes, v.2, 1.

Eignor, D., Taylor, C., Kirsch, I. & Jamieson, J. (1998). *Development of a Scale for Assessing the Level of Computer Familiarity of TOEFL Examinees* (ETS Research Report No. 98-60). Princeton, NJ: Educational Testing Service.

Elley, W. B. (1992). *How in the world do students read?*. The Hague: IEA.

Ethington, C.A. (1990), Gender Differences in mathematics : an international perspective, Journal for Research in mathematics Education, v.21, 1, 74-80.

Fan, X. (1997), Gender Differences in mathematics Achievement: findings from the National Education Longitudinal Study of 1988, Journal of Experimental Education, v.65, 3.

Farkas, G. (1996). *Human Capital or Cultural Capital? Ethnicity and Poverty Groups in an Urban School District.* New York: Aldine de Gruyter.

Fennema, E. (1996). Mathematics, Gender and Research. In: Towards Gender Equity in mathematics Education. Dordrecht/Boston/London: Kluwer Academic Publishers, pp. 9-26.

Feijgin, (1995) -> Thematic Report Gary

Fierros, E.G. (1999), Examining Gender Differences in mathematics Achievement on the Third International mathematics and Science Study, Paper presented at the annual meeting of the AERA (ED 431602).

Fisher, C.W., and Berliner, D.C., eds. (1985). *Perspectives On Instructional Time*. New York: Longman.

Fogarty, G., Cretchley, P., Harman, C., Ellerton, and Konki, N. (2001). Validation of a Questionnaire to Measure mathematics Confidence, Computer Confidence, and Attitudes towards the Use of Technology for Learning mathematics. Math*ematics Education Research Journal*, Vol. 13, No. 2, 154-160.

Forgasz, H. J., Leder, G. C, And Gardner, P. L. (1999) The Fennema-Sherman mathematics as a Male Domain scale reexamined. *Journal for Research in mathematics Education, 30*(3), 342-348

Friedrich, H.F. (1995). Analyse und Förderung kognitiver Lernstrategien. *Empirische Pädagogik*, 9, 2, 115-155.

Fredrick, W.C., and Walberg, H.J. (1980). Learning as a Function of Time. *Journal Of Educational Research* 73: 183- 194.

Gamoran, G., & Berends, M. (1987). The effects of stratification in secondary schools: Synthesis of survey and ethnographic research. *Review of Educational Research,* 57, 415-435.

Ganzach, (2000) -> Thematic Report Gary

Ganzeboom, H.B.G., de Graaf, P.M., and Treiman, D.J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, *21*, 1-56.

Geissler, J. And Horridge, P. (1993). *University students' computer knowledge and commitment to learning*. Journal of Research on Computing in Education. 25, 347-365.

Gettinger, M. (1984).Achievement as a Function of Time Spent in Learning and Time Needed for Learning. *American Educational Research Journal*, 21: 617-628.

Grisay, A. (1997). *Évolution des acquis cognitifs et socio-affectifs des élèves au cours des années de collège.* Les Dossiers d'Éducation et Formation 88, Paris: Ministère de l'Éducation Nationale, Direction de l'Évaluation et de la Prospective,

Grissom, J.G. & Shepard, L.A. (1989). Repeating and dropping out of school. In Shepard, L.A. and Smith, M.L. (Eds.), *Flunking grades: Research and policies on retention.* Philadelphia, PA: Falmer, 34-63.

Grouws, D.A. and Cebulla, K. J. (2000). *Improving student achievement in mathematics*. Educational Practices Series, UNESCO, International Academy of Education/International Bureau of Education.

Guida, F.V.; Ludlow, L.H.; & Wilson, M. (1985) The Mediating Effect of Time-On-Task on the Academic Anxiety/Achievement Interaction: A Structural Model. *Journal of Research and development in Education*, 19: 21-26.

Guthrie, J. and Wigfield, A. (1997). Reading Engagement. Motivating Readers through Integrated Instruction. Newark (Del.): International Reading Association.

Hacket, G. and Betz, N. (1989). An Exploration of the mathematics Efficacy/mathematics Performance Correspondence. *Journal of Research in mathematics*, Vol. 20, 3, 261-273.

Hamilton, S. F. (1990). *Apprenticeship for Adulthood: Preparing Youth for the Future*. New York: Free Press.

Hamilton S. and Hurrelmann, K. (1993). Auf der Suche nach dem besten Modell für den Übergang von der Schule in den Beruf - ein amerikanisch-deutscher Vergleich. *Zeitschrift für Sozialisationsforschung und Erziehungssoziologie, 13,* 194-207.

Hanushek, E.A. (1994). Money might matter somewhere: A response to Hedges, Laine, and Greenwald. *Educational Researcher, 23*, 4, 5-8.

Harnishfeger, A. & Wiley, D. E.: The teacher learning process in elementary schools: Asynoptic of view. *Curriculum Inquiry* 6 (1975), S. 4-43.

Heathers, G. (1969). Grouping. In: R. Ebels (ed.). *Encyclopedia of Educational Research*. New York: Macmillan.

Hedges, L.V., Laine, R.D., Greenwald, R. (1994a). Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes. *Educational Researcher, 23*, 3, 5-14.

Hedges, L.V., Laine, R.D., Greenwald, R. (1994b). Money does matter somewhere: a reply to Hanushek. *Educational Researcher, 23*, 4, 9-10.

Helmke, A. & Weinert, F. E.: *Bedingungsfaktoren schulischer Leistungen*. In: Weinert, F. E. (Hrsg.): Psychologie des Unterrichts und der Schule (Enzyklopädie Psychologie, Serie Pädagogische Psychologie, Band 3), Göttingen 1997, S. 71-176.

Heynemann, S. P., and Loxley, W. A. (1982). Influences on Academic Achievement across High and Low Income Countries: A re-analysis of IEA Data. *Sociology of Education*, Vol. 55, No. 1, 13-21.

Heynemann, S. P., and Loxley, W. A. (1983). The Effect of Primary-School Quality on Academic Achievement across Twenty-Nine High- and Low-Income Countries. *American Journal of Sociology*, Vol. 88, No. 6, 1162-94.

Hicks, M. (1989). *The TOEFL computerized replacement test: Adaptive conventional measurement* (ETS Research Report No. 89-12). Princeton, NJ: Educational Testing Service.

Hossler, C.; Stage, F. & Gallagher, K. (1988). *The Relationship of Increased Instructional Time to Student Achievement*, Policy Bulletin: Consortium on Educational Policy Studies 1.

Howerton, D. L., Enger, J. M., and Cobbs, C. R. (1993). *Parental Verbal Interaction and Academic Achievement for At-Risk Adolescent Black Males.* Paper presented at the annual meeting of the Mid-South Educational Research Association, November 10, New Orleans, LA.

Huitt, W. (1995). *A systems model of the teaching/learning process*. Valdosta, GA: College of Education, Valdosta State University.

International Labour Organisation (1990). *International Standard Classification of Occupations: ISCO-88*. Geneva: International Labour Office.

Jencks, C. (1972). *Inequality: A reassessment of the effect of family and schooling.* New York: Basic Books.

Johnson, William L. And Others(1991), *A Factorial and Rasch Analysis of the Charles F. Kettering Ltd. School Climate Profile*. Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, April 3-7, 1991) (ED335369)

Kirsch, I., Jamieson, J., Taylor, C. & Eignor, D. (1998). *Computer Familiarity Among TOEFL Examinees* (ETS Research Report No. 98-59). Princeton, NJ: Educational Testing Service.

Kershaw, T. (1992). The effects of educational tracking on the social mobility of african americans. *Journal of Black Studies*, 23,1.

Kulik, C.C. & Kulik, J.A. (1982). Effects of ability grouping on secondary school students: a meta-analysis of research findings*. American Educational Research Journal, 19*, 415-428.

Kulik, J. A. and Kulik, C.-L. C. (1992). Meta-analytic findings on grouping programs. *Gifted Child Quarterly*, 36, 73-77.

Kranzler, J. H. and Pajares, F. (1997). An Exploratory Factor Analysis of the mathematics Self-Efficacy Scale-Revised (MSES=R). *Measurement and Evaluation in Counseling and Development*, Vol. 29, 4, 215-228.

Laguarda, K., Hightower, A., Leighton, M., Weiner, L. (1995) *Raising the Educational Achievement of Secondary School Students*, Washington,D.C.: Policy Studies Associates, Inc.

Lamont, M. and Lareau, A. (1988). Cultural capital: allusions, gaps and glissandos in recent theoretical developments. *Sociological Theory* 6: 153-68.

Lee, V.E. (2001). *Restructuring High Schools for Equity and Excellence: What Works.* Sociology of Education Series. New York: Teachers College Press.

Lee, V. E. June, 1993. Educational Choice: The Stratifying Effects of Selecting Schools and Courses. *Educational Policy*. 7(2),125-148.

Lee, V. E. and A. S. Bryk. April, 1988. Curriculum Tracking as Mediating the Social Distribution of High School Achievement. *Sociology of Education*. 61: 78-94.

Lee, V.E., Bryk, A.S., Smith, J.B. (1993) The organization of effective secondary schools. *Review of Research in Education, 19*, 171-268.

Lee, V.E., and J.B. Smith (1996), *High School Size: Which Works Best, and for Whom?* Paper presented at the annual meeting of the American Educational Research Association, New York, April 1996. ED 396 888.

Lehmann, R. (1996). Reading Literacy Among Immigrant Students in the United States and Former West Germany. In: Binkley, M, Rust, K. and Williams, T. (Eds.). *Reading Literacy in an International Perspective*. Washington D.C.: NCES.

Lepper, M. R. (1988). Motivational considerations in the study of instruction. *Cognition and Instruction*, 5, 289-309.

Levin, T. And Gordon, C. (1989). *Effect of gender and computer experience on attitudes toward computers.* Journal of Educational Computer Research, 5, 69-88.

Levine, D.U., & Lezotte, L.W.. (1990). *Unusually effective schools. A review and analysis of research and practice.* Madison, WI: National Center for Effercive School research and Improvement.

Loveless, T.B. (1998) The Tracking and Ability Grouping Debate, Thomas B. Fordham Foundation. http://www.fordhamfoundation.org/library/tracking.html.

Karweit, N. (1976). A Reanalysis of the Effect of Quantity of Schooling on Achievement. *Sociology Of Education* 49 (1976): 236-246.

Ma, X. (1995), Gender Differences in mathematics Achievement between Canadian and Asian Education Systems, Journal of Educational Research, v.89, 2, 118-127.

Ma, X. (1999). A Meta-Analysis of the Relationship Between Anxiety Toward mathematics and Achievement in mathematics. *Journal for Research in Mathematics Education*, 30, 5, 520-540.

McLanahan, S.S. and Sandefur, G. D. (1994). *Growing Up with a Single Parent: What Hurts, What Helps.* Cambridge, MA: Harvard University Press.

Mcleod, D.B. (1992). Research on affect in mathematics learning: A reconceptualization. In D. A. Grouws (ed.). Handbook of research on mathematics teaching and learning (pp. 575-596). New York: Macmillan.

Mare, R. D. (1981). Change and stability in educational stratification. *American Sociological Review* 46, 72-87.

Marjoribanks, K. (1997). Children of Single-parent Families. In: Saha, L. J. (Ed.). International Encyclopedia of the Sociology of Education (pp. 589-95). Oxford/New York/Tokyo: Elsevier.

Marsh, H. W. (1990). Self Description Questionnaire (SDQ) II: A theoretical and empirical basis for the measurement of multiple dimensions of adolescent self-concept: An interim test manual and a research monograph. San Antonio, TX: The Psychological Corporation.

Marsh, H. (1999). *Evaluation of the Big-Two-Factor Theory of Motivation Orientation: Higher-order Factor Models and Age-related Changes*. Paper presented at the 31.62 Symposium Multiple Dimensions of Academic Self-Concept, Frames of Reference, Transitions, and International Perspectives: Studies From the SELF Research Centre. Sydney: University of Western Sydney.

Marsh, H. W. and Hau, K. (2002). Cross-cultural Generalisability of the Internal-External Frame of Reference Model: Relations Between Academic Self-Concept and Achievement in 26 Countries.

Mayeske et al. (1972). *A Study of Our Nation's Schools*. Washington. D.C.: US Department of Health, Education and Welfare.

Meece, J., Wigfield, A. and Eccles, J. (1990). Predictors of maths Anxiety and Its Influence on Young Adolescents' Course Enrollment and Performance in mathematics. *Journal of Educational Psychology*. Vol. 82, 1, 60-70.

Meier, D.W. (1996), The Big Benefits of Smallness, *Educational Leadership*, 54, 1 (September 1996): 12-15.

Meredith, D, Russell, N., Blackwood, L, Thomas, J. And Wise, P. (1999). *Real Time. Computers, Change and Schooling*. Canberra: Department of Education, Training, and Youth Affairs.

Middleton, J,.A. and Spanias, P.A. (1999). Findings, Generalizations, and Criticisms of the Research. *Journal for Research in mathematics Education*, Vol. 30, 1, 65-88.

Monk, D.H. (1992). Education productivity research: An update and assessment of its role in education finance reform. *Education Evaluation and Policy Analysis, 14*, 307-332.

Mortimore , P., Sammons, P., Stoll, L., Lewis, D. and Ecob, R. (1988). *School Matters: The Junior Years. Somerset.* Open Books (reprinted in 1995 by Paul Chapman/London).

Mullis, I.V.S., Martin, M O., Gonzalez, E J., Gregory, K D., Garden, R A., O'Connor, K M., Chrostowski, S J., and Smith, T A. (2000). *TIMSS 1999 international mathematics report : findings from IEA's repeat of the Third International mathematics and Science Study at the eighth grade.* Chestnut Hill MA : Boston College. International Study Center, 2000.

Multon, K. D., Brown, S. D. and Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counselling Psychology*, 38, 30-38.

Murphy, J. (1990). Principal instructional leadership. In: Thurstone, P. and Lotto, L. (Eds.). *Advances in Educational Leadership* (pp. 163-200). Greenwich, CT: JAI Press.

Nakamura, J. (1988). Optimal Experiences and the Use of Talent. In: M. Ciskszentmihalyi and I.S. Ciskszentmihalyi (Eds.). *Optimal Experience: Psychological studies of flow in consciousness*. New York: Cambridge University Press, pp. 150-171.

Oakes, J. (1985). *Keeping track. How schools structure inequality.* New Haven: Yale University Press.

O'Neil, H.F. and Herl, H. E. (1998). *Reliability and validity of a trait measure of self-regulations.* UCLA/CRESST AERA.

Organisation for Economic Co-operation and Development. (1999). *Classifying Educational Programmes. Manual for ISCED-97 Implementation in OECD Countries.* Paris: OECD Publications.

Organisation for Economic Co-operation and Development. (2001). *Knowledge and Skills for Life: First Results from PISA 2000*. Paris: OECD Publications.

Owens L. and Barnes J., (1992) *Learning Preference Scales*, Hawthorn: Australian Council for Educational Research.

Phelps, M. S. and Prock, G. A. (1991). Equality of Educational Opportunity in Rural America. In: DeYoung, A. J. (Ed.). *Rural Education. Issues and Practice* (pp. 269-312). New York/London: Garland Publishing.

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1993). Reliability and redictive validity of the Motivated Strategies for Learning Questionnaire (MLSQ). *Educational and Psychological Measurement* 53, 801-813.

Pelgrum, W., Janssen R, I. & Plomp, T. (1993). *Schools, teachers, students, and computers: A cross-national perspective.* Enschede: University of Twente, Center for Applied Educational Research. (ERIC No. ED 372 734)

Peklaj, C. & Vodopivec, B. (1999). Effects of cooperative versus individualistic learning on cognitive, affective, metacognitive and social processes in students. *European Journal of Psychology of Education*, 14(3), pp. 359-373.

Postlethwaite, N. and Ross, K. (1992). *Effective schools in reading. Implications for educational planners*. The Hague: IEA.

Purkey, S.C. and Smith, M.S. (1983) Effective schools: a review. Elementary School Journal, 83, 427-452.

Persell, C. et.al. (1992). Differential asset conversion: class and gender pathways to selective colleges. *Sociology of Education*, 65, 3, 208-25.

Postlethwaite, T.N. (Ed.) (1995). *The international encyclopedia of national systems of education. (2nd edition)*. Oxford: Pergamon.

Ramseier, E. (1997). Individuelle Determinanten der mathematikleistung. In: U. Moser, E. Ramseier, C. Keller and M. Huber (1997), *Schule auf dem Prüfstand*. Zürich: Ruegger, 114-136.

Raywid, M.A. (1996), *Taking Stock: The Movement to Create Mini-Schools, Schools-Within-Schools, and Separate Small Schools*. Madison, Wisconsin: Center on Organization and Restructuring of Schools. ED 393 958.

Rheinberg, F. (2001). Bezugsnormen und schulische Leistungsbeurteilung. In: Weinert, F. E. (Ed.). *Leistungsmessungen in Schulen* . Weinheim: Beltz, 59-71.

Roderick, M. (1993). The Path to Dropping Out: Evidence for Intervention. Westport, CN: Auburn House.

Rossi, R.J., and Stringfield, S.C.(1995). *Education Reform and Students At Risk. Studies of Education Reform.* Palo Alto, California: American Institutes for Research in the Behavioral Sciences; and Baltimore, Maryland: Center for Research on the Education of Students Placed At Risk.

Rutter, M., Maughan, B., Mortimore, P., Ouston, J, and Smith, A. (1979). *Fifteen Thousand Hours. Secondary schools and their effects on children.* Cambridge, M.A.: Harvard University Press.

Saha, L. J. (1997). Introduction: The Centrality of the Family in Educational Processes. In: Saha, L. J. (Ed.). International Encyclopedia of the Sociology of Education (pp. 587-88). Oxford/New York/Tokyo: Elsevier.

Sammons, P., Hillman, J, and Mortimore, P. (1995). *Key characteristics of effective schools: a review of school effectiveness research*. London: London University Institute of Education, ISEIC for OFSTED.

Seegers, G. and Boekaerts, M. (1996). Gender-related Differences in Self-Referenced Cognitions in Relation to mathematics. *Journal for Research on mathematics Education*. Vol. 27, No. 2, 215-240.

Selwyn, N. (1997). *Assessing Students' Ability To Use Computers: Theoretical Considerations For practical research.* British Educational Research Journal, 23, 1, 47-59.

Schaffer, E. C., Nesselrodt, P. S., and Stringfield, S. (1994). The Contribution of Classroom Observation to School Effectiveness Research. In: Reynolds et. al. (Eds.). *Advances in School Effectiveness Research and Practice* (pp. 111-132). Oxford/New York/Tokyo: Pergamon.

Scheerens, J. & Bosker, R. J. (1997) *The foundations of educational effectiveness*. Oxford, New York and Tokyo, Pergamon Press.

Schildkamp-Kundiger, E., ed. (1982), An international Review of Gender and mathematics,(ED 222326).

Schmidt, W.H. *et al*. (1996). *Characterizing Pedagogical Flow: An investigation of mathematics and Science Teaching in Six Countries*. Kluwer, Dordecht.

Schulz, W. (2002) Constructing and Validating the Questionnaire composites. In R.J. Adams and M.L. Wu. *PISA 2000 Technical Report*. OECD Publications. Paris

Schunk, D. H. (1990). *Socialization and the Development of Self-Regulated Learning: The Role of Attributions.* Paper presented at the Annual Meeting of the American Educational Research Association, Boston, MA, April 16-20.

Seltzer, J. (1994). Consequences of marital dissolution for children. *Annual Review of Sociology*, 20, 235-66.

Sewell, W. H. and Hauser, R. M. (1975). *Education, Occupation and Earnings: Achievement in the Early Career.* New York: Academic.

Siniscalco, M. T. and Ross, K. N. (1997). *The establishment of an International Reading Resources Scale: An exploratory study using Modern Item Response Theory.* Paris: UNESCO.

Slavin, R. E. (1983). When does cooperative learning increase student achievement? *Psychological Bulletin*, 94(3), 429-445.

Slavin, R. E. (1990a). Ability grouping in secondary schools: A response to Hallinan. *Review of Educational Research*, 60, 505-507.

Slavin, R. E. (1990b). Achievement effects of ability grouping in secondary schools: A best evidence synthesis. *Review of Educational Research*, 60, 471-499.

Slavin, R. (1990). On making a difference. *Educational Researcher*, 19(13), 30-44.

Slavin, R. E. (1994). Quality, appropriateness, incentive, and time: a model of instructional effectiveness. *International Journal of Educational Research, 21*(2), 141-158.

Slavin, R. (1996). *Education for All.* Contexts of Learning series. Lisse, The Netherlands: Sweits & Zeitlinger.

Skinner, E. A., Chapman, M. and Baltes, P. B. (1988). Control, means-ends, and agency-beliefs: A new conceptualization and its measurement during childhood. *Journal of Personality and Social Psychology*, 54, 117-133.

Skinner, E.A., Wellborn, J. and Connell, J. (1990). What It Takes to Do Well in School and Whether I've Got It: A Process Model of Perceived Control and Engagement in School. *Journal of Educational Psychology*. Vol. 82, 1, 22-31.

Stigler, J.W.; Gonzales, P.; Kawanaka, T.; Knoll, S. and Seerano, A. (1999). *The TIMSS Videotape Classroom Study. Methods and Findings from an Exploratory Research Project on Eight-grade mathematics Instruction in Germany, Japan, and the United States*. Washington, D.C.: Us Department of Education, Office of Educational research and Improvement, NCES 99-074.

Stringfield, S..; Ross, S.M..; Smith, L., Eds. (1996). *Bold Plans for School Restructuring: The New American Schools Designs*. Washington, DC: Office of Educational Research and Improvement (ED), Caroll, J.B.: A model of school learning. In: Teacher College Record 1963 (64. Jg.), S. 723-733.

Stroh, L (1991), High School Student Evaluation of Student Teachers: How Do They Compare with Professionals? *Illinois School Research and Development*; v27 n2 p81-92 Win 1991 (EJ432373).

Tanner, H. & Jones, S. (2000). *Becoming a Successful Teacher in mathematics*. London/ New York: Routledge.

Tate, W. F. (1997). Race-Ethnicity, SES, Gender, and Language Proficiency Trends in mathematics -Achievement: An Update. *Journal for Research on mathematics Education*. Vol. 28, No. 6, 652-679.

Teachman, J. (1987). Family background, educational resources and educational attainment. *American Sociological Review* 52, 548-57.

Teddlie, C. (1994a). The Study of Context in School Effects Research: History, Methods, Results, and Theoretical Implications. In: Reynolds et. al. (Eds.). *Advances in School Effectiveness Research and Practice* (pp. 73-84). Oxford/New York/Tokyo: Pergamon.

Teddlie, C. (1994b). The Integration of Classroom and School Process Data in School Effectiveness Research. In: Reynolds et. al. (Eds.). *Advances in School Effectiveness Research and Practice* (pp. 85-110). Oxford/New York/Tokyo: Pergamon.

Teddlie and Reynolds 2000

Werf, M. P. C. van der (1995). *The Educational Priority Policy in the Netherlands: Content, Implementation and Outcomes.* The Hague: SVO.

Van de Werfhorst, H. (2001), Field of Study and social Inequity. Four types of Educational resources in the process of Stratification in the Netherlands, PHD thesis, Nuffield College, Oxford.

White, K. R. (1982).The relation between socioeconomic status and educational achievement. *Psychological Bulletin*, 91(3), 461-81.

Wigfield, A., Eccles, J. S. & Rodriguez, D. (1998). The development of children's motivation in school contexts. In P. D. Pearson. & A. Iran-Nejad (Eds.), *Review of research in education* (Vol. 23, 73-118). Washington DC: American Educational Research Association.

Wilkerson, David J.; Manatt, Richard P.; Rogers, Mary Ann; Maughan, Ron (2000), Validation of Student, Principal, and Self-Ratings in 360 Degree Feedback (registered) for Teacher Evaluation, *Journal of Personnel Evaluation in Education*; v14 n2 p179-92 Jun 2000 (EJ619020)

Williams, T., Long, M., Carpenter, P., and Hayden, M. (1993). *Entering Higher Education in the 1980s*. Canberra: AGPS.

Willms, J. D. (1987). Patterns of Academic Achievement in Public and Private Schools: Implications for Public Policy and Future. In: Haertel, E. H., James, T., and Levin, H. M. (Eds.). *Comparing Public and Private Schools. Volume 2. School Achievement* (pp. 113-134). New York/Philadelphia/London: The Falmer Press.

Zimmerman, B.J. (1999). Commentary: toward a cyclically interactive view of self-regulated learning. International Journal of Educational Research, 31, 545-551.

Zimmermann, B. J. and Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology*, 82, 51-59.

Zimmerman, B.J. & Schunk, D.H., & (Eds). (1989). Self-Regulated Learning and Academic Achievement. Theory, Research and Practice. New York: Springer-Verlag. From Teaching to Self-Reflective Practice. New York: Guilford Press.

CURRENTLY UNCITED

*Doyle, W. (1985). Effective secondary classroom practices. In M.J. Kyle (Ed.) *Reaching for Excellence: an effective school sourcebook.* Washington, D.C.: US Government Printing Office.

*Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (in press). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist*. (Paper presented at the AERA Annual Meeting 2000).

*Pintrich, P. R., & DeGroot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82, 1, 33-40.

*Schunk, Dale H., & Zimmerman, Barry J. (Eds). (1998). Self-Regulated Learning: From Teaching to Self-Reflective Practice. New York: Guilford Press

*Slavin, R. E. (1999). Comprehensive Approaches to Cooperative Learning. *Theory into Practice*, 38 (2), p74-79.

*Taylor, B. M., Pressley, M., & Pearson, D. (2000). *Effective Teachers and Schools: Trends across Recent Studies.* National Education Association.

*Marsh, H. (1990). Influences of Internal and External Frames of Reference on the Formation of math and English Self-Concepts. *Journal of Educational Psychology*, V82, 1, 107-116.

*Adam Gamoran, A., Nystrand, M., Berends, M. and LePore, P.C. (1995) An Organizational Analysis of the Effects of Ability Grouping, A study of the effects of ability grouping in 92 honors, regular, and remedial English classes in eighth and ninth grade, *American Educational Research Journal*, Winter 1995.

*Braddock, J.H. & McPartland, J.M. (1990). *Alternatives to tracking*. Baltimore, MD: Johns Hopkins University, Center for Research on the Effective Schooling of Disadvantaged Students.

*Lee, V. E., & Bryk, A. S. (1989). A Multilevel Model of the Social Distribution of High School Achievement. *Sociology of Education, 62*, 172-192.

*McPartland, J.M. (1990). Staffing decisions in the middle grades: Balancing quality instruction and teacher/student relations. *Phi Delta Kappan* 71 , 465-469.

*Marsh, H. W., Shavelson, R. J. & Byrne, B. M. (1992). A multidimensional, hierarchical self-concept. In R. P. Lipka & T. M. Brinthaupt (Eds.), Studying the self: Perspectives across the life-span.

*Nuttall, D.L., Goldstein, H., Prosser, R., & Rasbash, J. (1989). Differential school effectiveness. *International Journal of Educational Research, 13*, 769-776.

*Rettig, M.D. & Canady, R.L. (1999). The effects of block scheduling, *School Administrator*, 56, 3, 14-20

*FORGASZ, H., LEDER, G., and KAUR, B. (2001) Who can do maths- boys/girls? Beliefs of Australian and Singaporean Secondary School. *Asia Pacific Journal of Education, 21*(2), 106-116

*HARNISCH, D.L. AND RYAN, K.E. (1983), An investigation of the relationship of achievement motivation with achievement in mathematics for the students in the United States and Japan, Paper presented at the annual meeting of the MWERA (ED 235886).

*HOLLAND, J. L, (1985) Making Vocational Choices, A theory of vocational personalities and work environments, Odessa, Florida, PAR.

*MURNANE, R. J, AND RAIZEN, S.E. Eds, (1988), Improving indicators of the Quality of Science and mathematics Education in grades K-12. Washington: National Academy Press.

*West, Sandra S.; Denton, Jon J. , (1991), An Empirical Validation of the Instrument: Student Perceptions of Teaching Effectiveness. Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, April 3-7, 1991).(ED330724)

*Peterson, Kenneth D.; Wahlquist, Christine; Bone, Kathie (2000), Student Surveys for School Teacher Evaluation. *Journal of Personnel Evaluation in Education*; v14 n2 p135-53 Jun 2000 (EJ619019)

*Gamoran, A. (1987). The stratification of high school learning opportunities. *Sociology of Education*, 60, 135-155.

*Gamoran, A., and Mare, R. D. (1989). Secondary school tracking and educational inequality: Compensation, reinforcement, or neutrality. *American Journal of Sociology*, 94, 146-183.

*Gamoran, A. December, 1992. The Variable Effects of High School Tracking. *American Sociological Review.* 57: 812-828.

*Hansen, N.M. (1997), Social and Economical Inequality in the Educational Career. *European Sociological Review*, 13, 305-321.

*François Rastoldo et Bernard Favre (dir.) (2000), *Hétérogénéité et différenciation au cycle d'orientation. Classes hétérogènes et classes à sections au 7e degré : Carrières d'élèves et discours d'acteurs.* Genève: Service de la recherche en éducation, Décembre 2000.

*Weber, J. M. (1988). Dropout rates for students in different high school curricula. *Journal of Vocational Education Research*, 13, 35-49.

*Shavit, Y. and Blossfeld, H.P. (1993) Persistent Inequality. Changing Educational Attainment in Thirteen Countries. Boulder, CO: Westview Press.

ANNEX 1: CRITERIA FOR SELECTING THEMES

184. It is not only important to identifying possible policy-relevant themes but also to choose wisely from among the many possibilities so that the strengths of the PISA design are capitalised upon and PISA's contributions to policy-makers' and educators' needs is maximised.

185. This document uses a set of criteria that have been established by the INES Network A for the selection of themes:

- First, they must be of enduring policy relevance and interest. That is, a theme should have *policy relevance*, capture policy-makers' attention, address their needs for data about the performance of their educational systems, be timely, and focus on what improves or explains the outcomes of education. Further a theme should be of *interest to the public*, since it is this public to which educators and policy-makers are accountable.
- Second, the themes must provide an internationally comparative perspective and promise significant added value to what can be accomplished through national evaluation and analysis. This implies that themes need to be both relevant (i.e., of importance) and valid (i.e., of similar meaning) across countries.
- Third, there must be some consistency in approach and themes with PISA 2000.
- Fourth, it must be technically feasible and appropriate to address the issues within the context of the PISA design. That is, the collection of data about a subject must be *technically feasible* in terms of methodological rigor and the time and costs (including opportunity costs) associated with data collection.

186. These criteria are explained in the following in more detail.

Policy relevance

187. The call for tender for PISA 2003 notes the following policy issues for analytic work:

> *Because mathematics is the major domain assessed in the second PISA survey cycle, ... PISA [should] portray important aspects of learning and instruction in the field of mathematics, including the impact of learning and teaching strategies on achievement as well as the impact of school organisation and structures in promoting active student engagement with learning. The questionnaires should also address issues related to: mathematics efficacy and students' engagement, motivation and confidence with mathematics; mathematics and gender; and students' planned educational pathways. The quality of the school's human and material resources; issues of public/private control, management and funding; school level information on the instructional context and institutional structures are important aspects as well.'*

188. When working through these policy issues, it is useful to consider three dimensions:

- *Quality of educational outcomes and educational provision* (e.g.: does a country have qualified teachers and well-equipped schools? Are children taught a challenging curriculum, with appropriate teaching methods, in classrooms with positive social and disciplinary climate, optimal use of school time, supportive and motivated teachers? Does the young generation develop the knowledge, skills, values, learning habits and interests required to live in their society?)
- *Equality of educational outcomes and equity in educational opportunities* (e.g: are there significant disparities in the way human and material resources are distributed across a country's schools? Is the curriculum equally well implemented and equally demanding everywhere? To what extent does the distribution of certain characteristics of the school intake (such as gender, socio-economic background, immigrant status, etc.) reflect possible segregation trends in a school system? How large is the overall variation in cognitive and

affective outcomes among students? To what extent can these disparities be attributed to differences in the quality of instruction delivered by the schools attended?

- *Adequacy and effectiveness of resource management* (e.g. to what extent are school resources used / the curriculum taught / teaching methods implemented / outcomes obtained, in ways which are. appropriate to needs? Does the educational system allow for adaptation to regional, local or individual needs? What aspects should be considered for the system to become more cost-effective?)

189. Since trend indicators are a new important component in PISA 2003, each of the above three issues can be considered both in terms of the *current situation* — how good, equitable and adequate is the school system currently in a particular country, compared to other countries? — and in terms of *trends* — did the quality, equity and adequacy of a particular school system improve or deteriorate during the last years, compared to other countries?

190. In the design of PISA the view was taken that the context questionnaires will contain a set of *core data*, which are to be included in every cycle. Such data will provide a 'monitoring function' and will consist of basic demographic variables and others deemed to be of 'high-value.' Examples of demographic variables include gender, grade level, and socio-economic status (such as items on parents' occupation and education). Examples of 'high-value' variables may include, for instance, first language or language spoken at home.

191. Regular inclusion of items such as these permit comparisons over time and further analyses of themes to be addressed with data from PISA 2000. Other monitoring variables such as students' ethnicity, school location may be better suited to national questionnaires (as, for instance, variables like race and ethnicity currently are treated). Such items may not be universally relevant or of interest to all participating countries and therefore should *not* be a part of the core data or context questionnaires.

192. In addition PISA's context questionnaires will also address themes of particular interest to policy-makers and relate specifically to the primary substantive domain of the cycle.

Cross-national comparability

193. The fact that PISA is an international study adds a number of important benefits, but it also causes some constraints.

194. Questionnaire instruments used in national assessments can usually provide most of the information needed to help answering questions like those given above (particularly the adequacy and effectiveness issues), in ways that are tailored to the specific characteristics of the system. For instance, they are more appropriate than international surveys to inform educational authorities on whether all aspects of a new curriculum were implemented effectively in their country's schools, or on the proportion of students that met specific competency standards in national exams, or on possible local negative effects of an otherwise beneficial nation-wide innovation, etc.

195. However, national assessments are unable to provide two very important pieces of information that can only be derived from regular international assessments like PISA:

- Only international assessments can inform national authorities about the extent to which other school systems 'do better' than their system, in terms of student outcomes, instructional delivery, teacher's qualification or professional development, etc., or about whether the school organisation in other countries results in fewer disparities in the quality of instruction delivered and in a lower impact of social background on student's outcomes; and
- Only international assessments can show whether the evolution over time of some indicator was positive (or negative) in their country compared to other countries.

196. This unique potential of international assessments implies that priority should be given, in the design of the PISA 2003 Questionnaires, to aspects for which comparisons across time and space are both possible and of real value.

197. The constraints come in the need for the use of questions and coding schemes that retain their meaning and validity in an international context. The PISA 2000 questionnaires made important advances in this area. For education level, PISA used the International Standard Classification of Education (ISCED; OECD, 1999b). The use of this international coding scheme allowed for a common coding of the educational programmes that cater for 15-year-olds. For occupational status PISA 2000 used the International Socio-economic Index (ISEI) developed by Ganzeboom, de Graaf and Treiman (1992). For a range of other variables derived especially for use in PISA cross-national validation analyses were undertaken and are reported in the PISA technical report (Schulz, 2002).

198. There are, however, limits to the extent to which standardised coding schemes or items with cross-national validity can be constructed. For example, there are still considerable challenges to be faced in the construction of indices of cultural and social capital.

Building on PISA 2000

199. Griffith (2000) stressed in her evaluation of PISA 2000 the importance of taking a long term perspective on PISA. She wrote that it is:

> *'critically important to reach closure on the issues that will be emphasised in the background questionnaires and the body of research that will be produced by [each] cycle of PISA so that the design of the study (including the sample, questionnaires, and analyses) can be made on an a priori basis.'*

200. The benefit of doing so, she adds, will be that PISA's design will intentionally fit its goals.

201. While this activity has not yet been undertaken several historical factors affect the number and range of policy themes that can be included in PISA 2003 and, thus, can be addressed in the context questionnaires. These factors, which reflect the preferences and previous decisions of participating countries and PISA's overall purposes, include:

- the need to include high-quality items with high explanatory value (such as family education and students' socio-economic status) as the core in each cycle of PISA; and
- an interest in comparisons over time and the need to retain, through identical items in the context questionnaires, the ability to make these comparisons and to establish trends.

Technical feasibility

202. A strong focus on literacy in key domains, an age-based population definition that is particularly adequate to inform on the *yield* of the educational systems, and a cyclical design are the three major characteristics of the PISA assessment that will provide policy-makers with an ongoing means for monitoring how well students are prepared for participation in society, lifelong learning, and personal fulfilment.

203. These three characteristics, however, also impose some constraints on what can be asked of the PISA studies. In its current form, PISA is a 'yield-oriented' assessment. That is, its features are optimal for the purpose of accurately summarising the outcomes levels of 15-year-olds. The PISA design is not optimal for other purposes. In particular, with an age-based population and a literacy orientation, the very concept of 'instructional context' cannot be well defined (in terms of the wide range of educational settings where the 15-year-old students are to be found, and the range of settings in which their literacy will be developed). This means that PISA has a relatively limited descriptive and/or explanatory power for aspects of the

instructional context that are not 'universal' dimensions, shared by all 15-year-old students. For instance, while PISA can certainly provide information on the impact on achievement of variations in 'global' context characteristics (such as the disciplinary climate of the school attended, or of the level of perceived supportiveness of teachers), it would be much less likely that the effects of more class-specific or grade-specific dimensions such as the effects of the curriculum implemented or of the instructional methods used be properly explored and give policy meaningful results. A careful examination of factors of this type would require additional components to be included in the PISA design (such as an age-grade based design, use of intact classes, and/or longitudinal analyses) which the BPC may wish to consider for future survey cycles.

204. The yield orientation of PISA and the broad definition of literacy that underlies the development of all assessment instruments also contribute in limiting the information that can be collected on the effects on achievement of specific curricula or particular methods of instruction. The student's performance as measured in PISA is the result of years of exposure to a multitude of learning situations, both in school and out of the school. Therefore, it is very unlikely that the characteristics of the instruction received during the particular school year when the student is tested affected in a significant way the skills, knowledge or processes assessed. Again, this means that the context variables of potential interest in PISA are likely to be mainly found among the 'global', pervasive and long-lasting characteristics of the school environment (such as the socio-economic intake of the school, the aspirations of the pair group, the type of track attended or the general level of demand of the curriculum taught), rather than among more contingent characteristics such as the qualification of the current mathematics teacher or his/her teaching practices.

205. Perhaps the best way to express some of the limitations of the PISA design is in the coverage of the framework shown above that can be attained.

206. PISA does not include activities that directly focus on collecting data regarding relevant national level antecedents and contexts, that is cells 8 and 12. A range of such data is however available from the OECD education indicators programme.

207. Because PISA currently does not survey teachers there are limits on the availability and relevance of data on some classroom contexts and antecedents, such as teacher characteristics and qualifications, and on classroom processes such as pedagogical practices and curriculum content (Cells 6 and 10). Any information on these aspects must be collected from either students or at the school level. Similarly any perspectives on students' behaviours and such must be obtained from the students themselves.

208. The classroom variables that will be of most utility in PISA will likely be, as mentioned above, *global* variables that are common to most students in a given setting and appear as relatively stable characteristics of the educational context. Examples might include: perception of the supportiveness of their teachers, perceptions of the teaching styles used, perceptions of the disciplinary climate in the school, and so on. Since the information will be collected from students it is expected that those aspects in the teaching practices that are sufficiently 'prominent' or 'visible' to the student - the aspects that are likely to have a real and direct impact on them - will be emphasised. Fortunately, there is literature indicating that secondary school students have reasonably accurate perceptions about their teachers' behaviour and the teaching styles they are exposed to (Johnson *et al.*, 1991; Stroh, 1991; Wilkerson *et al.*, 2000).

209. Similarly, at the school level (cells 7 and 11), PISA will likely be best served by focussing on questions that are related to relatively broad and 'stable' features such as school type, school structure, school resources, school climate, school management, most of which are known to have some impact on student's achievement, according to the school effectiveness literature.