International Education Journal Vol 5, No 3, 2004
http://iej.cjb.net

Monitoring the Learning and Teaching of Science in a Changing World

J.P. Keeves
Flinders University, School of Education *john.keeves@flinders.edu.au*

This article addresses the issues involved in monitoring the teaching and learning of science in a changing world. It examines the development of cross-cultural studies of educational achievement, particularly in the field of science, including the theoretical basis of the studies and the models advanced and used in evaluation and more recently in the monitoring of change over time. In addition ten specific issues are identified for investigation into the critical problems facing learning and teaching of science across the world at the beginning of the twenty-first century with particular reference to the PISA studies being conducted by the Organisation for Economic Cooperation and Development.

Science teaching, science learning, cross-national studies of educational achievement, monitoring of educational achievement, models of science learning and teaching

INTRODUCTION

The six decades since the cessation of hostilities in World War II have been accompanied by remarkable changes in the teaching of science in most countries of the world. The developments that occurred prior to and during the years of war, with respect to the electronic computer, the atomic bomb, and telecommunications, together with the discovery of the double helix shortly after the war, and the writings of Conant (1947) '*On Understanding Science*' and Popper (1959) '*The Logic of Scientific Discovery*' led to a major reform in science education in the late 1950s and early 1960s that spread across the world. Not only was logical positivism largely rejected and biology and earth science introduced into the school curriculum, but the teaching of physics and chemistry were also substantially changed together with the introduction of science into the primary school curriculum, as well as the teaching of science as a compulsory subject at the lower and middle secondary school levels. At the upper secondary school level, policies of raising the school leaving age and staying on at school beyond the upper age of compulsory schooling led to a marked rise in the popularity of biology, particularly among girls. In addition, ideas associated with the relations between science, technology and society and the history of scientific thought were introduced into the science curriculum, sometimes as minor subjects, and more frequently intertwined within the teaching of specific topics.

However, developments in the field of computer science and the emergence of information and communications technology that became divorced from the main fields of science in the school curriculum have served to draw students away from the subjects of physics and chemistry at both the upper secondary and higher education levels. As a consequence, in most countries of the world there is today a desperate shortage of younger teachers of physics and chemistry to replace an aging teacher work force. Moreover, the teaching of science as a core subject in most countries of

the world throughout the primary school years and up to the age of 15 years or the end of compulsory schooling is, during the first decade of the twenty-first century, struggling to capture the interest and imagination of students, particularly in the fields of physics and chemistry.

At the same time as the science curricula were being reformed in the mid-1960s, educational research emerged as a new and vigorous field of inquiry, making use of new technology and the power of electronic computers to conduct cross-national surveys of educational achievement in the subject areas of mathematics in 1964, and in science in 1970-71. These studies have been followed intermittently by repeat surveys to monitor change in the teaching of these subjects across the years of schooling. The international and comparative nature of these survey studies have revealed a surprising uniformity in the teaching of these two subjects, initially across the Western world and more recently in the Asia Pacific, Eastern Europe, Latin America and Middle East regions. The teaching of these subjects in schools would seem to have developed initially from Germany and Great Britain, and subsequently from the United States, France and Russia to influence the curricular content of these subjects across the world. While there are some variations in the structure of curricula, there is little variation in the content being taught, except in Mathematics in Francophone countries, in Physical Geography and Earth Science in Eastern European countries, and in Behavioural Science in countries linked to Russia. Since many of the surveys undertaken have sought to analyse curricular differences, rather than to emphasise comparisons with respect to levels of achievement, the studies have probably served to unify the teaching of these subjects around the world, rather than to increase diversity.

These cross-national studies were initiated in the late 1950s by a group of prominent educational research workers who met in England and at the UNESCO Institute of Education in Hamburg to discuss common problems in the conduct of educational research. From their deliberations they recognised the need for a comparative research program that was empirically oriented and that investigated problems that were common to many national systems of education. They saw the world of education as a natural laboratory in which different countries were experimenting with different strategies of teaching and learning. By examining the naturally occurring differences between countries in both the conditions of learning and educational outcomes, they argued that it might be possible to identify significant factors that influenced the outcomes of education. Consequently, they formed an organisation in 1961 known as the International Association for the Evaluation of Educational Achievement, now commonly referred to as IEA, to develop a program of research that would be both comparative and cooperative in order to pursue their objectives. The teams of scholars who guided the development of the IEA research studies worked under the guidance of Torsten Husén and Neville Postlethwaite, who were able to attract contributions from leading scholars from Europe and the United States initially under the auspices of the UNESCO Institute for Education in Hamburg, Germany, and subsequently the Institute of International Education at the University of Stockholm in Sweden, and currently, an administrative centre at The Hägue in The Netherlands. Thus, the PISA 2006 testing program draws on a body of educational research particularly in the areas of mathematics and science that has been carried out cross-nationally for almost 40 years. While the PISA 2006 program is focused primarily on science at the 15-year-old or middle secondary school levels, rather than the 14-year-old and terminal secondary school stages, that were investigated by IEA it relates to the final year of compulsory schooling in many of the countries involved and thus provides information with respect to a cohort of students, that is not affected significantly by dropping out from school. Where formerly the 14-year-old age level served a similar purpose, the 15-year-old level is, in 2006, more appropriate. However, present day studies draw extensively on the thinking carried out during earlier investigations.

THE THEORETICAL BASIS OF EARLIER STUDIES

In the field of science education earlier studies were reported by Comber and Keeves (1973), Keeves (1992a and b), Postlethwaite and Wiley (1992), and Rosier and Keeves (1991) and in mathematics by Husén (1967), Robitaille and Garden (1989) and Rosier (1980). These publications in the main do not provide accounts of the lengthy discussions that underpinned the design and development of the investigations that were undertaken. However, leading scholars from Europe and the United States in the fields of education and the social and behavioural sciences contributed to these discussions with working papers, that are now stored in archives in California in the United States. All that can be done in this article is to summarise some of the key ideas that influenced the planning of these early investigations.

No grand theory was advanced to provide a framework for the systematic study of education on a world basis, although Holmes (1981) had sought to provide one. Nevertheless, investigations that have been undertaken for the evaluation of educational achievement would seem to have had their origins in the work directed by Tyler for the Eight-Year Study (Aikin, 1942). These ideas were elaborated initially by Tyler (1949), and subsequently by Bloom et al. (1956), Bloom, Krathwohl and Masia (1964), Bloom, Hastings and Madaus (1971) and revisited by Tyler (1986) in a largely American context. While Bloom was deeply involved in the founding of the IEA movement, the IEA studies were developed through the involvement of scholars from all parts of the world, and these American based publications do not adequately represent the richness of the views of those who shaped these investigations.

The outcome of these scholarly discussions has led to the formulation of a series of models that have been employed in the cross-national studies conducted by IEA with respect to the following problem situations: (a) curriculum implementation, (b) time and school learning, (c) causal models of school learning, (d) cross-national models of educational achievement in a national economy, (e) an input-output-utilisation model, (f) a retentivity model for school learning beyond the years of compulsory schooling, and (g) an educational environment model for the investigation of the influence of the environments of the home, the classroom and the peer group on educational achievement. Each of these models is described briefly in the section that follows.

MODELS FOR THE EVALUATION OF EDUCATIONAL ACHIEVEMENT

It is important to note that the models under consideration were, in the main, developed within the International Association for the Evaluation of Educational Achievement (IEA) by teams of scholars, rather than particular individuals.

Curriculum Implementation Theory

From the Gränna Workshop conducted by IEA in Sweden in 1971, which examined in detail the seminal work '*The Handbook of Formative and Summative Evaluation of Student Learning'* by Bloom, Hastings and Madaus (1971), came the model (see Figure 1) of curriculum implementation that has been tested, in part, in reporting the results of the First and Second IEA Science Studies and the Second IEA Mathematics Study (Keeves, 1974; Keeves, 1992a; Postlethwaite and Wiley, 1992; Robitaille and Garden, 1989; and Rosier and Keeves, 1991).

The curriculum can be considered to exist at three levels: (a) the intended curriculum, (b) the implemented curriculum, and (c) the achieved curriculum, which are influenced by the antecedent and the contextual factors operating at the systemic, classroom and student levels respectively. The **intended curriculum** is usually specified by political bodies and authorities in charge of an education system. However, in some systems the responsibility to specify what is taught resides with the board of an individual school, or with each individual teacher within a school. The **implemented curriculum** is the second level in the curriculum sequence. It is the task of each

individual teacher to interpret the intended curriculum by translating it into a set of specific learning experiences that are considered appropriate for the particular group of students in a class. The **achieved curriculum** is the third stage. It refers to the extent to which individual students have learnt from the experiences that were planned and organised for them. Figure 1 shows that the intended curriculum is set in the context of the education system; the implemented curriculum is located in the context of the school or classroom; and the achieved curriculum relates to the individual student.

An important aspect of the implemented curriculum involves the opportunity that the students under survey had to learn specific content topics from the larger pool of knowledge that is considered both necessary and desirable knowledge for teaching to particular age and grade groups. Three aspects of curriculum validity were identified by Rosier and Keeves (1991), namely:

(a) to what extent does a particular intended curriculum match the more general curriculum formed by the body of content that might be taught;

(b) to what extent do the test items cover the intended curriculum; and

(c) to what extent do the test items relate to what is taught in the intended curriculum.

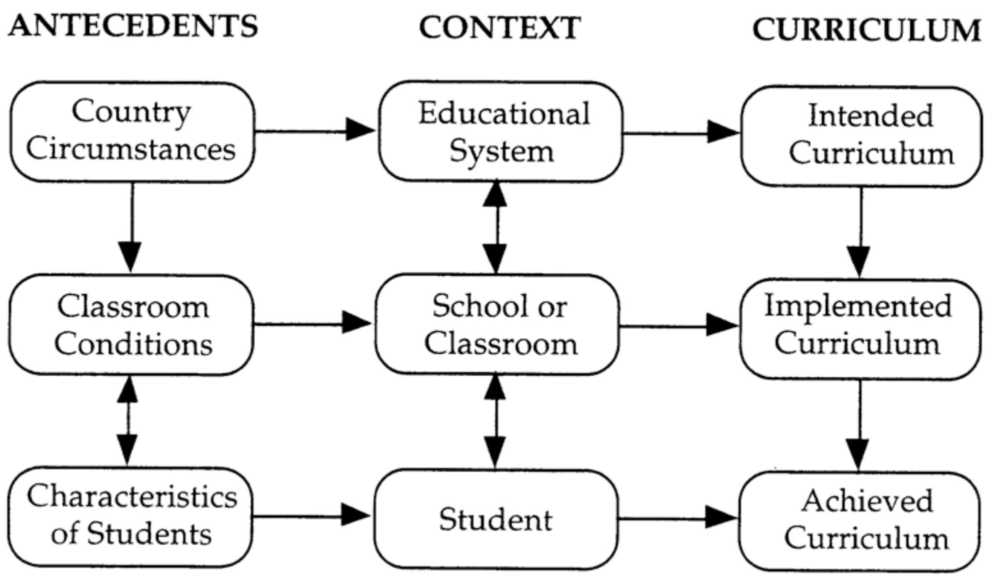

Figure 1. The context and components of the school curriculum

It is clear that the implemented curriculum is dependent on the intended curriculum, and the achieved curriculum upon the curriculum that is implemented in the classroom.

Carroll's Model of School Learning

The model of school learning advanced by Carroll (1963) has been the source of many theoretical discussions of the factors influencing educational achievement. The IEA studies provided opportunities for the investigation of this theory and for the derived models to be tested empirically (Carroll, 1975). Carroll developed this model in order to investigate the prediction of success on complex learning tasks. Three variables were specified in terms of time: (a) **aptitude** that involved the amount of time a student would require to learn a task to a specified criterion, given motivation, opportunity to learn and optimal quality of instruction; (b) **perseverance** that involved the amount of time which a student was willing to engage in active learning; (c) **opportunity to learn** that involved the amount of time provided for learning in a specific program. In addition there were two further variables that were not specified in terms of time; (d)

ability to understand instruction that was dependent on the quality of instruction provided; and (e) **quality of instruction** that involved the structuring of the learning task, the effectiveness of presentation and the skills of the instructor. These two additional variables interacted with each other and with the three variables directly involving time. Bloom (1974) drew attention to the relevance of IEA findings on **time** for the study of school learning and gradually the significance of time as a key concept in learning has come to be accepted.

A Causal Model of School Learning

In the planning of the Second IEA Science Study a causal model derived from Carroll's model of school learning was advanced in 1981 which was subsequently tested in the analyses of the data collected in that study at the 10-year-old, 14-year-old and terminal secondary school levels (Keeves 1992a). This causal model has also been examined in detail in several doctoral theses. The model of performance in science is shown in Figure 2 and its dependence on Carroll's ideas is immediately evident. An important aspect of this model is the identification of the levels of operation and analysis of the variables included. Such a model not only guides the selection of information collected and the variables constructed, but also guides the analysis of data.

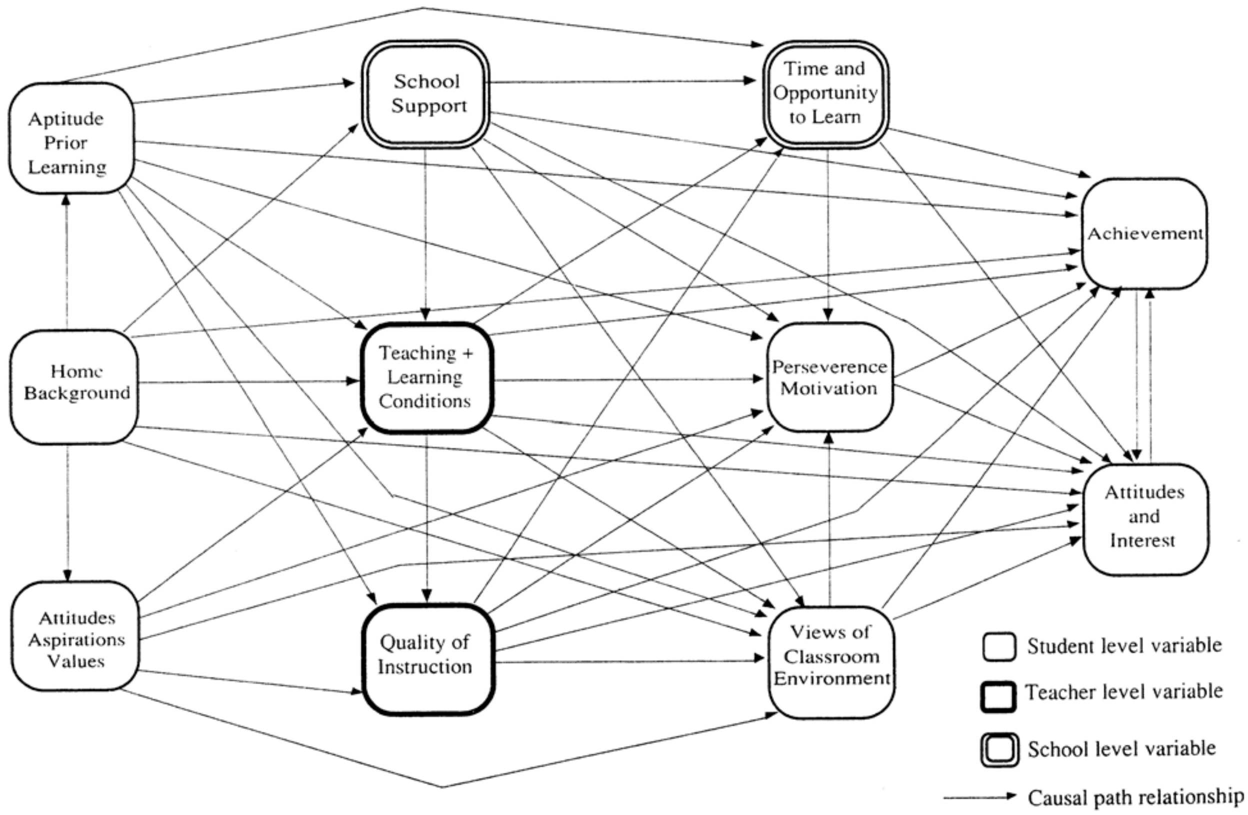

Figure 2. A model of student performance

A Cross-national Model of Educational Achievement in a National Economy

In 1967, during the planning phase for the IEA Six Subject Study, a conference was held at Lake Mohonk in the United States, which sought to develop a 'cross-national model of educational achievement in a national economy'. A paper by Dahlöff (1967) developed a scheme for the **educational process** that applied in cross-national settings, that is shown diagrammatically in Figure 3.

Of particular interest for policy making in education are the frame variables. However, they depend on (a) the environment and economy, (b) demand for manpower, (c) curriculum content, and (d) the objectives of education.

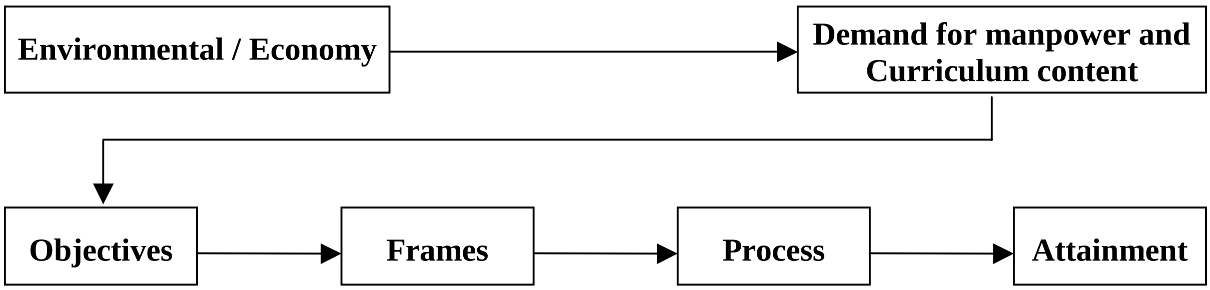

Figure 3. A cross-national model of educational achievement in a national economy

Input-Output-Utilisation Model of Education

The Lake Mohonk Conference also advanced an Input-Output-Utilisation model of education which is presented in Figure 4 (Super, 1967) that included many significant components, namely: (a) financial (b) production conditions, (c) structure and operations (educational structure, equipment, agents, curriculum, and instructional methods), (d) outputs (knowledge, skills, attitudes, participation, attainment level), and (e) utilisation (employment, community involvement, and family activity). While causal influences, student movement and financial flow were taken into consideration, insufficient thought was given at that stage to the testing of such a complex model. The task remains for more specific hypotheses and multilevel causal models to be advanced that are amenable to testing, and build on the ideas advanced by Super (1967). The model shown in Figure 4 extends Super's original model through the addition of (f) social and cultural capital and (g) key aspects of national development, including human development.

Retentivity Model for School Learning

The basic idea on which this model is built is that in each country there is the same underlying distribution of intellectual ability in the complete age cohort and that differences in mean scores and variances in any cross-national test at a stage beyond the age of compulsory schooling arise as a result of the selection procedures that operate within the country. This may seem a gross over-simplification of a complex situation, and if the model does not fit the data then the model must be rejected. However, in situations where this model has been tested it would seem to provide an adequate account of the situation.

The simple assumptions involved are:

(a) the scores in each country would be normally distributed over the whole age group, if all persons in the age group had taken the tests;

(b) these hypothetical distributions are the same for all countries; and

(c) the populations under survey in each country involve the most able persons in the age group in each country.

Using these assumptions the expected mean scores and variances can be calculated for the groups forming the selected proportions of the age group for each country which can be compared with the observed mean scores and variances (Walker, in Husén, 1967, pp.135-139).

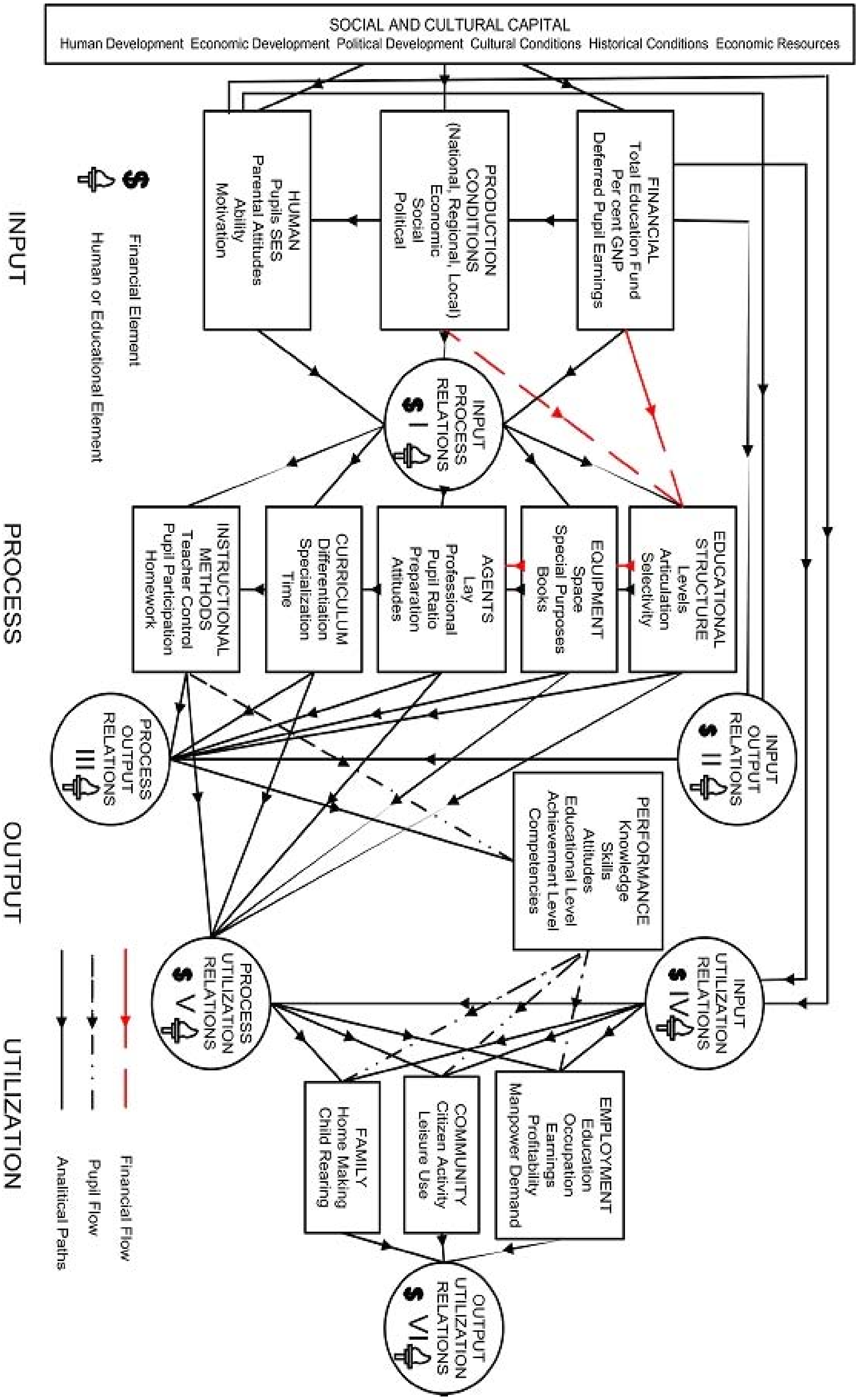

Figure 4. A functional process organic model of education based on Super (1967)

The model employed for the estimation of the mathematical achievement test scores and their variance when a proportion of the total age group participates in the study is shown diagrammatically in Figure 5, from which the formulae given below are derived.

$$Mean = \frac{y}{q}$$

$$variance = \frac{1 - y}{q\left\{\left(\frac{y}{q}\right) - k\right\}}$$

Where q = proportion selected,

y = ordinate of normal curve at point of cut-off,

k = point of cut off.

If there is a correlation between the variable operating in selection from the population and the achievement test score under survey, then the above formulae for the mean and variance are:

$$Mean = \frac{ry}{q}$$

$$variance = r^2\left(\frac{y}{q}\right)\left\{\left(\frac{y}{q}\right) - k\right\}$$

Where r is the correlation.

Figure 5A shows the normal distribution for mathematics scores with a cutting score *k* to select proportion *q* of the age cohort, assuming that those students selected would obtain the highest scores. Since this is not likely to be the case in a real situation the model is modified, as is shown in Figure 5B with a correlation *r* between the variable operating to select from the population, that is assumed to be normally distributed, and the mathematics scores obtained.

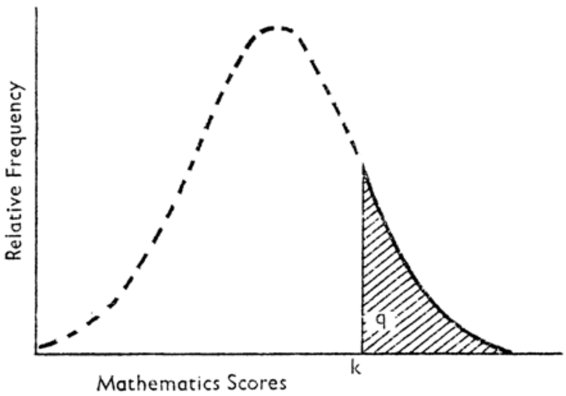

Figure 5A. Hypothetical distribution of achievement test scores for age cohort with cutting score k to select proportion q of the cohort

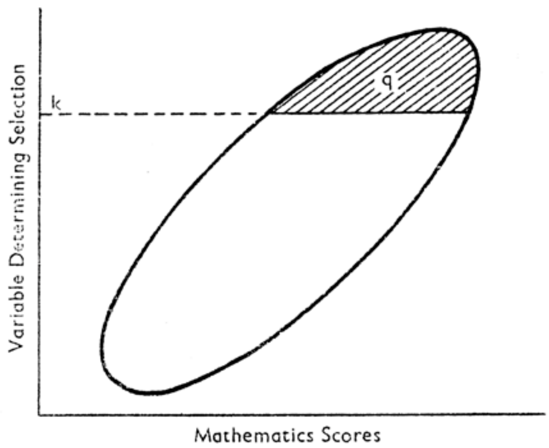

Figure 5B. The correlation surface relating achievement test scores to the selection variable, with cutting score k forming proportion q of the cohort

Educational Environment Model for Educational Achievement

A model was also developed from unpublished working papers that were discussed at the Lake Mohonk Conference and from discussions with Bloom in the Department of Education at the University of Chicago. This model was advanced for investigating the factors that influenced the change in performance of classroom groups of students over time. It was hypothesised that the environments of the home, the classroom and the peer group operated to influence the rate of change in achievement over time. Moreover, it was argued that these environments were best characterised by structural, attitudinal and process dimensions that could be specified in each educational setting within a society. The interrelations between the three environmental influences for each of the three dimensions could then be estimated from a recursive path model in which prior achievement was permitted to influence certain aspects of the home environment, which in turn was permitted to influence the classroom and the peer group environments. The model is shown in Figure 6 (Keeves, 1972, pp. 38-40).

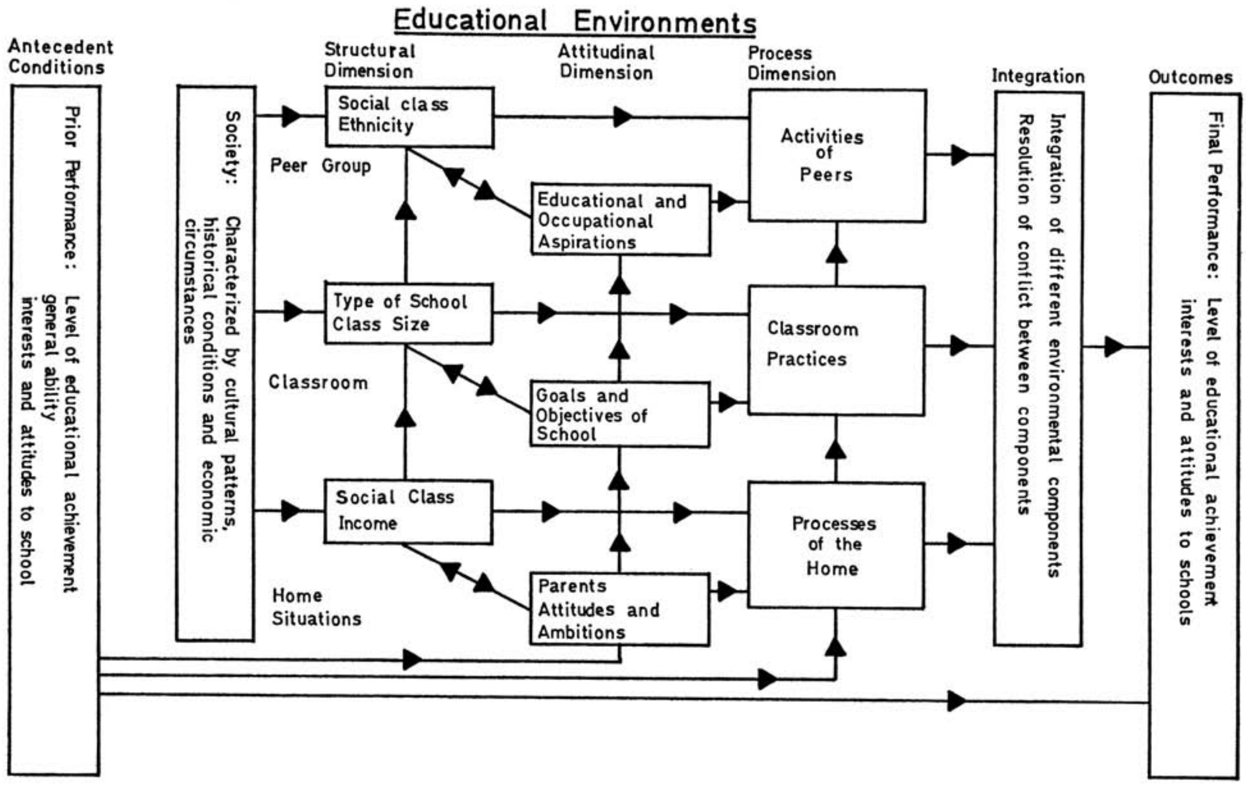

Figure 6. Model for the Study of Educational Environments.

OVERVIEW OF MODELS

The increased elegance of the methods of analysis that can be used in future studies that provide for both continuous and categorical data and the estimation of path models at two levels permit the testing of models involving both student and classroom data. Moreover, the advances made in the field of multilevel modelling for the estimation of cross-level interactions associated with three level models permit the analysis of data at the student, school and country levels, with provision for country-school interactions, or country-student interactions, or school-student interactions. In addition, with knowledge of cross-level interactions that might be expected, it is now possible to estimate models that provide for more than three levels, namely student, class, school and country levels, as well as estimating cross-level country specific effects.

Simplistic analyses that merely involve the estimation of bivariate effects at a single level are commonly highly suspect when the real situation is considerably more complex. Furthermore, studies that fail to control for spurious effects or report simple relationships with aggregated data are also very misleading, when such effects are mediated through other variables, or where aggregation bias has severely distorted the nature and magnitude of an effect. Consequently the models advanced in the past must today be reconstructed with consideration given to the analytical procedures that are being employed in the estimation of the parameters of a model and the testing of a model for fit to observed data and thus the adequacy of the model. Moreover, it should be noted that the computer programs that can be employed are progressively being developed to remove restrictions being imposed on the models being tested. Thus it would be inappropriate at the present time to specify the computer programs that could be used in any particular situation because the emergence of a new version of a program could change completely the comments that might be made.

Nevertheless, the models presented in the paragraphs above are essentially cross-sectional models and need to be developed and extended to examine effects when data that are longitudinal at the country, school or student levels over more than two occasions are under examination. Consequently, in a series of survey studies that involves monitoring at regular intervals over time, new models must be developed that permit not only the investigation of change in the criterion measures, but also in the effects of rates of change in the predictor variables. Fortunately, this is a field of data analysis that is under development at the present time, and new and powerful computer programs are being produced. The opportunity to monitor change over time and to investigate the factors influencing change give rise to exciting developments in the field.

TEN ISSUES FOR INVESTIGATION

In this paper I have deliberately chosen to identify ten issues for investigation, that I see to involve critical problems which are currently largely ignored in much of what I read and hear in the fields of science education that are influenced by current educational perspectives in the Western world of Europe and the United States. Moreover, in discussing these ten issues I have chosen to list them in a decreasing order of relevance and perceived importance, although I recognise that those issues lower on the list are dependent on issues or themes higher on the list. Furthermore, I see opportunities for reorienting the nature of the problems investigated when data from monitoring and longitudinal studies become available.

Issue 1: Mathematics and Science Achievement and the Labour Market

Elaboration. Scholars in the countries of East Asia are to some degree very puzzled by the results that have been produced by the series of studies undertaken by IEA over the past 40 years and more recently by PISA that provide consistent evidence of the superior average level of performance by students in a group of countries in East Asia in comparison with the countries of Europe and North America in the fields of mathematics and science (see, Spearitt, 2003). Hanushek and Kimko (2000) have addressed the issue of whether mathematical and scientific knowledge and skills are a major component of the human capital that is relevant to the quality of the labour force of a country. They have used performance in cross-national testing programs of mathematics and science achievement as an indicator of quality of human capital and have estimated its relationship to measures of economic performance. Labour force quality differences measured in this way have been shown to have strong relationships with growth rates in national economies. These effects were particularly strong for the countries of the Asia-Pacific region (Tuijnman, 2003, pp. 1088-9). This contrast between the East and West countries is possibly grounded in the differences between the approaches of science educators to reform the science and technology curricula as portrayed in an article in the *International Handbook of Educational*

Research in the Asia Pacific Region by Kok-Aun Toh and Ngoh-Kehang Goh (2003, pp. 1243-1256). Alternatively, it may be related to differences in approaches to learning between countries in the East and the West (Watkins, 2003, pp. 449-453).

Implementation. At least three aspects of this issue can be addressed by the undertaking of analyses of PISA data. First, there is the possibility of replicating the work of Hamushek and Kimko (2000) with data collected from IEA studies Second, there would appear to be the need to obtain information from the developers of curricula with respect to curriculum reform in the fields of mathematics, science and technology, and the role of constructivist thinking in curriculum design in contrast to modelling approaches that serve to accelerate the cognitive development of students and sustain an interest in mathematics, science, and technology as proposed by Shayer and Adey (1981, 2002). The third aspect would involve an examination of the expectations for post-school employment and training of the samples of students in each county that are related to science and technology based skills, since these are the fields in many countries where there is a significant shortage of scientifically and technologically skilled personnel. A further aspect would require the development of a measure of generalised skill to indicate the competence of the students to use the Mathematics and Scientific knowledge in work and societal situations and to relate this measure to rates of growth in productivity.

Issue 2: Science and Technology Curricula in Schools

Elaboration. The introduction of information and communications technology into schools during the past two decades has been largely divorced from the teaching of science and mathematics, in so far as science and mathematics teachers have been reluctant to be concerned with issues that involve the application of their disciplinary knowledge to technological problems. As a consequence, the focus of technology teaching has been solely on ICT, and often merely on the development of word processing skills, the use of spread-sheets and the internet, without recognition of the many ways in which electronic computers are transforming the conduct of mathematical and scientific inquiry or the ways in which computer based technology is being applied in industry and commerce. The time has come for mathematics, science and technology teachers to reform their curricula in order to integrate the use of technology not only into the processes of teaching and learning, but also into the processes of doing mathematics and science, with recognition that the dynamic power of computerised technology is capable of changing the modes of thinking and learning by students in schools, in working life and in lifelong learning and personal development.

Implementation. A case must be argued most strongly for the introduction of a theme that is concerned not with the use of ICT in the science classroom, but with the use of technology in the fields of mathematics, and science and its uses in the home, the learning environment and the workplace in its many different forms. This leads to the consideration of such uses as (a) individualising instruction, (b) diagnostic testing with immediate feedback and informed correctives, (c) reflective thinking in learning with different modes of immediate verification, (d) the use of experimental design that involves the testing of models and hypotheses by simulation, (e) inductive thinking through control of the experiences provided, (f) enrichment experiences through direct access to video disc presentations, (g) the collection, storage and presentation of data, (h) searching the internet for information, (i) computing where complex calculation is required, and (j) robotic performances of complex tasks.

Issue 3: Further Education and Career Choice in Science and Technology

Elaboration. Issues 1 and 2 raise problems that are related indirectly with a choice of a career in the fields of science and technology, but without considering directly the factors that might influence the possibility of such a choice among students of 15 years of age. The choice of a

career in a field of endeavour that is related to science and technology is influenced by many factors that operate at the individual, classroom, school and systemic levels. Consequently it would be incomplete and inadequate merely to examine the issue of career choice in terms of individual student characteristics, when the home and the peer group as well as the climates of the classroom and the school are also involved. Moreover, it is evident in some countries that there is widespread acceptance that it is cheaper to recruit a highly skilled and scientifically and technologically trained workforce from other countries than to attempt to lift the level of education in a country that has a pressing need to recruit such highly skilled workers.

Implementation: In order to examine the effects of factors at different levels both within and between countries it is necessary to obtain from students, parents, teachers, and school administrators information that would help to address the issues of career choice in science and technology. Of particular importance would be the information that could be readily obtained from students on their expected occupation in a form that would enable the occupation to be coded as scientifically oriented, technologically oriented, or as a non-scientific and non-technological occupation. It should be noted that 12 years ago a study that would investigate such issues through secondary data analysis had to be abandoned, because the computer programs that would permit an analysis at the student, school and systemic levels were not at that time available. This situation has been remedied and such a study is clearly warranted. Moreover, planning should be undertaken from the outset to conduct, if possible, a longitudinal study that followed individuals from the age of 15 years over a decade or more. The identification of factors that influenced the career paths of such students as they moved from school to further education and on into stable careers that were scientifically and technologically oriented or otherwise would be of considerable value.

Issue 4: Attitudes to Science and their Effects

Elaboration: The issue of career choice, raised in the preceding section, depends at least in part, on the question of the contribution of attitudes towards science and technology in decision making by individual students for a career in these fields. The investigation of attitudes towards science and technology is not without its problems and there are skeptics who challenge the meaningfulness of the information obtained through the administration of questionnaires with Likert type scales. Nevertheless, considerable progress has been made towards the measurement of attitudes and values using item response theory to obtain interval scaled data, that permits not only the use of attitudinal data as predictor measures, but also the use of attitudinal variables as criterion, or mediating variables in both multilevel and multivariate models to build a greater understanding of the different factors that operate to influence educational outcomes.

Implementation: It is important that a range of attitude scales should be constructed to access a range of attitudinal dimensions which contribute towards achievement and behavioural outcomes. Attitude scales that have been employed in past studies include: (a) beneficial effects of science; (b) interest in learning science; (c) ease of learning science, and (d) career interest in science. (see Keeves, 1992 a and b). In addition, the classification of attitudes, interests, and orientation advanced by Klopfer (1971) provide a valuable framework for the assessment in the affective domain in the fields of science and technology. It is important to note that at least seven items should be employed within each scale in order to obtain scales with sufficient content coverage to provide a meaningful and internally consistent scale and so warrant the use of Rasch scaled scores. Too often in survey research studies, space and time constraints restrict the attitudinal information collected to single items. While such limited information has some value it would appear to have relatively little explanatory power that would lead to a greater understanding of the effects of attitudes on achievement and the choice of a career oriented towards science and technology.

Issue 5: Gender and Student Performance and Attitudes

Elaboration: Over the past 40 years cross-national comparative studies have helped trace and explain the differences in the changes that have occurred in the composition and performance by sex of student and teacher in different fields of education, as well as achieving a greater understanding of gender effects on educational outcomes. Several indexes associated with equality between the sexes in educational participation and provision have shown clearly identifiable changes over time, namely:

- substantial falls in the ratio of male to female students at the terminal year of secondary schooling, with similar falls at grade levels beyond the stage of compulsory schooling;
- falls in the ratio of male to female teachers at the secondary school and higher education levels;
- in countries where some single-sex schools formerly existed, there were substantial declines in the ratio of single-sex to coeducational schools (Keeves, 1992a and b).

These effects would appear to indicate a changing role of women in society and a pressure for greater equity in educational and occupational participation between the sexes.

These changes have also been accompanied by a reduction in the difference between the sexes in achievement in the physical sciences and mathematics, together with a greater degree of participation by girls in these subjects (Keeves 1992 a and b). Nevertheless, more than 20 years ago there were signs of disturbing effects emerging at the terminal secondary school level in some countries where the ratio of male to female students in academic schooling fell to as low as 0.5 (Keeves, 1992 a and b). It would appear that for various reasons boys in some countries were opting out from formal academic education, preferring to move more rapidly towards employment and the earning of money and away from the more rigorous subjects of mathematics and science. Baker and Jones (1993) have shown a clear relationship between the increasing rate of participation of women in the labour force and the rate of reduction of differences between boys and girls in mathematics achievement at the lower and middle secondary school levels. This study would seem to provide evidence for relationships between societal forces that influence participation in the labour force and educational outcomes in so far as differences between the sexes are involved.

Implementation: There is clearly a need to monitor change in the differences between the sexes in participation, achievement and attitudes in the field of mathematics, science and technology education over time, and to seek explanation between the effects observed and societal forces. The changes that have been observed and associated with gender effects indicate quite clearly the likelihood that such effects are not only societal in nature, but also run the risk of giving rise to effects that are prejudicial to the best interests of boys. An inquiry into these effects within one culture is unlikely to provide an understanding of their nature and consequences, and the opportunity to undertake a cross cultural investigation at a level of education where the whole of an age group is involved, namely at the last stage before differential dropping out from school can occur, is likely to be highly rewarding.

Issue 6: Students' Engagement in Learning Science and Technology

Elaboration. The most powerful factor and the most strongly contested in curriculum planning, but the most poorly researched and monitored in the practice of teaching and learning is that of time in its many different forms, including curricular time, engaged time, homework time, and time spent on different types of activity in the classroom. Carroll's (1963) model of school learning draws attention to the importance of time in teaching and learning in the classroom

situation, but the findings of the limited body of research carried out to test this model continue to be largely ignored. While there are considerable difficulties in measuring the different relevant aspects of time involved in school learning, this should not lead to a complete rejection of all attempts to estimate the effects of time on the learning of science and technology in schools, in the home, as a member of the peer group in hobby activities, and through informal learning by visiting museums, watching television programs that present scientific information, reading science magazines, searching the internet for scientific material, and engaging in bird watching, naturalist field studies and the systematic recording of weather data. The significance of the different aspects of time for learning in science and other fields warrants not only the examination of the explanatory power of such time based measures, but also the examination of factors that influence engagement and participation.

Implementation. The collection of information on the different relevant aspects of time involved in school learning presents problems for those responsible for the construction of questionnaires for the students, parents, teachers and school administrators, but the use of multilevel analytical procedures enable the effects of data that may provide under-estimates at the individual level to be aggregated and examined meaningfully at the school and national levels. Moreover, time is a measure that has the same meaning across different countries and cultures and as a consequence sound comparisons can be made at the higher levels of analysis. It is also important to recognise that the effects of time on learning are cumulative. Consequently, the strong evidence (Keeves, 1992 a and b; Postlethwaite and Wiley, 1992) of accumulative effects of learning at the primary school stage on learning at the secondary school stage should be taken into consideration in estimating the accumulative effects of curricular time. Furthermore, consideration should be given to the estimation of curricular time spent on learning in the different fields of science, Biology, Chemistry, Physics, Earth Science, Psychology and Physical Geography, as well as the estimation of viewing time and viewing participation by school-aged students of science-oriented television programs.

Issue 7: Parental Investment in Scientific and Technological Literacy

Elaboration. While the teaching profession has largely ensured that the fields of science and technology are in the main divorced from one another in the curricula of the schools, this is not so in the home, in leisure time activities and in the work place. Moreover, it should be clearly recognised that technological development is the major reason for the significant place that science now holds in the school curriculum at all levels. Consequently in any examination of the effects of parental investment in scientific literacy it would seem essential that this separation is not maintained. Parental investment involves both a time commitment and a financial commitment, since not only do children learn from their parents by example and joint participation, but they also learn as a consequence of the provision of books, scientifically-oriented toys and games, access to a computer and recording equipment, and opportunities to visit museums, technology display centres, zoological gardens and natural history excursions. Effects that are commonly ascribed to socio-economic status and parental educational level or to the resources of the home are, in general, more appropriately attributed to the practices of the home and the things that parents do to develop particular interests, attitudes and values in their children as well as to develop their children’s levels of cognitive and physical skills.

Implementation. Information on parental investment in scientific and technological literacy is probably best obtained from the parents themselves, although the administration of a questionnaire to parents is likely to give rise to substantial missing data, which is difficult to cover by imputation, since failure to respond is itself an indicator of lack of parental investment in their children’s education. Consequently, information on parental investment is necessarily obtained from the students under survey, thus limiting the amount and nature of the information that can be

collected. It should also be noted that information obtained from students or parents can be aggregated to the school and community levels, and would thus form a variant of 'social and cultural' capital. Perhaps, at the school and country level a term involving the idea of 'scientific and technological' capital is required in order to indicate the level of interest and involvement of each community in scientific and technological matters.

Issue 8: Science and the Environment

Elaboration. The issues of concern that relate to the environment are associated with the science and technology curricula of the schools in so far as some environmental problems have their origins in scientific and technological development. Moreover, commonly there is reliance on science and technology to monitor environmental change and to indicate whether environmental conditions are deteriorating. However, in some countries, Australia for example, concern for the environment is not associated with the science and technology curricula, but directly linked to learning about society and is thus seen to involve primarily attitude change rather than being based on knowledge and understandings. Consequently, the policy issues in the field of education involve not only the place that environmental issues hold in the science and technology curricula, but also the attitudes held by students concerning the beneficial and harmful effects of science, that arise from learning about the society in which the students live.

Implementation. The issues involving science and the environment require that not only should questions be asked in the cognitive tests that can be attributed to the structure and nature of the science curriculum, but questions should also be asked through the attitude questionnaires. In the assessment of attitudes it should be noted that evidence from cross-national surveys (Keeves, 1992a) has indicated that the concern for the beneficial aspects of science and concern for the harmful aspects of science do not lie along the same dimension and the use of separate scales would seem to be essential. However, responses to such scales may not arise from the structure and nature of the science and technology curricula but from other areas of the school curriculum, and information on the curricular source of such attitudes should also be sought. Furthermore, the sources of both information about environmental issues as well as attitudes also involve the home and the media as well as the peer group and the wider community. Moreover, it is largely through the home and the wider community that environmental degradation originates. From the educational viewpoint, not only are the levels of knowledge and attitude of importance, but the sources of ideas are of considerable importance if educational practices are to respond to the findings of the PISA study.

Issue 9: Teaching and Learning of Science

Elaboration. Science is the only subject in the school curriculum in which information obtained through the senses from the real world is used through the process of induction to form knowledge about the real world. In addition, when deductions are drawn from knowledge already held, science requires that the deduced ideas and relationships must be tested against observations from the real world, or through experiment and observation in the real world, before the ideas and relationships are accepted as an adequate account of those aspects of the real world under consideration. This latter process involves modelling the real world and the testing of the model against the real world. Some aspects of mathematics also involve a more extensive use of deductive processes, and statistics involves checking the adequacy of a model in a stochastic way. Learning through computer technology develops certain skills that are similar to learning in science, but the process of checking ideas and relationships in computerised learning involves immediate feedback from the machine in testing the ideas and relationships against observations drawn from the real world. While much of science teaching is expository and declarative, the processes of science teaching and learning largely involve the forming of models of the real world

through the operations of induction and deduction and the subsequent testing of these models. These processes are the bases of inquiry, and the critical operation of inquiry involves testing propositions, hypotheses or models, in the real world. No school subject, apart from science, seeks to build knowledge in this way. These operations place the learning of science in an almost unique position. While induction and the associated observations are important to science and are seen as part of the inquiry process, the key processes of modern science are the formulation of hypotheses and models, and the testing of the models is a critical part of the process. The deduction component and the design and conduct of an experiment is clearly investigatory in nature but the modelling process is central to learning about the tactics and strategies of science, and to understanding the nature of science, as well as building a deeper understanding of scientific principles, ideas and relationships.

The learning of science involves (a) training in observation and induction, (b) training in deduction to form hypotheses and models, (c) designing of experiments to test a model, and (d) conducting an experiment to test the proposed model. These processes are set in stages by Piagetian theory and Shayer and Adey (1981, 2002) argue that the learning of science not only involves the development of understanding of key scientific ideas and relationships, but also the development in schooling of those cognitive skills that are involved in scientific thinking at appropriate stages in the students' education.

Implementation. The administration of achievement tests in science at the 15 year-old age level is carried out at a stage when the student is in a transition phase of learning that involves the modelling process. As a consequence the test items employed should range across the SOLO Taxonomic levels described by Biggs and Collis (1992). These levels are not dissimilar to the levels proposed in the revised Bloom taxonomy by Anderson and his colleagues (Anderson, 2001).

Other aspects of the testing program in science and technology would seem to require: (a) an assessment of understanding the relationships between Science, Technology and Society, (Aikenhead and Ryan, 1992; Tedman, 2002), (b) attitudes associated with the learning of science discussed under Issue 4, and (c) an assessment of the classroom climate to indicate the emphases placed on inquiry and the modelling approach, in contrast to an over emphasis on factual knowledge and examination performance, or a concern for the interests and needs of individual students (see Keeves, 1992a).

Issue 10: Teachers of Science

Elaboration. The major problem in the learning and teaching of science in the countries of the Western world is the serious shortage of teachers in the fields of physics, chemistry and earth science. There would appear to be an adequate supply of teachers of biology, but with an over supply of female teachers. The problem of shortage of well qualified science teachers has existed in most Western countries for at least 15 years and is becoming more acute as older qualified teachers reach retirement age without an adequate supply of younger teachers graduating from universities to take their place. Because computer based technology is an emerging field there is a similar acute problem in this field that is accentuated because mathematics, the physical sciences and ICT teachers are drawn from the pool of graduating students who can find more lucrative work in the commercial and industrial ICT fields. There is no immediate solution to this problem, since the payment of salary supplements in order to attract sufficient numbers of teachers, who have been trained in mathematics, physical sciences and ICT, to fill the vacancies that exist, would exceed the financial resources available. Other alternatives would involve greatly increased class sizes, or would produce a marked imbalance between the regular teaching salaries and the salaries of those receiving salary supplements to teach mathematics, the physical sciences and technology.

Implications. The important information that does not currently appear to be available is baseline information to model the qualifications, salaries and fields of expertise of teachers currently in schools who are teaching in the different fields of the physical sciences, mathematics and ICT. Information on the supply of new teachers should be readily available from the training institutions, but the characteristics of the current teaching force in these fields, where there is known to be a shortage, would not seem to be available. As a consequence there would appear to be large numbers of mathematics and physical science and technology classes at the secondary school level in most countries of the Western world in which students are being taught by unqualified or under-qualified teachers. During recent decades the teaching of science in many countries has been extended throughout the primary school years where the teachers have little knowledge of the science required to undertake the effective teaching of scientific ideas. However, this critical problem while having implications for the PISA studies is outside PISA fields of survey, and OECD should investigate this problem through other studies that would give rise to models of the demand and supply of manpower in this field.

CONCLUSION

The nature of the PISA studies that involve the administration of testing programs at regular intervals of three and nine years for students at the terminal stage of compulsory schooling in most Western countries, does not commit PISA to seeking to provide a greater understanding of the educational processes that operate across the 12 years of schooling in most countries. IEA has for more than 40 years sought to investigate these educative processes in a wide range of school subjects at several levels of schooling using the countries taking part in IEA studies to form a natural laboratory. The role of the PISA studies would appear to be the monitoring of change over time in achievement and the educational processes that influence achievement. Consequently, from the ten issues discussed above, themes should be chosen that are within PISA's mandate of monitoring change. However, PISA is necessarily involved in providing, where possible, some explanation for the changes that it records and presents. This requires not only a knowledge of the educational processes that have been discussed in IEA studies, but also the modelling and testing of models that provide explanation for the changes that are observed in PISA studies. Scholars who are involved in the design and planning of PISA studies must accept this long range perspective of monitoring change, and must not only try to ensure that appropriate information on student achievement is collected, but also that appropriate information on explanatory variables is also assembled for use on later occasions. From the monitoring of change and the efforts to explain change it is likely that a deeper understanding of educational processes will emerge as Baker and Jones (1993) and Hanushek and Kimko (2000) have shown for problems that appeared unsolvable a decade or more earlier. It is fortunate indeed that analytical procedures were developed during the 1990s and that their development is continuing which permit the examination of multilevel longitudinal data. Moreover, the statistical information currently being assembled both rigorously and systematically by OECD (1992, et seq.) in the *Education at a Glance* series of publications provides data at the national level that was previously unavailable. However, the PISA studies that are conducted cross-nationally must be supported by intra-national longitudinal studies that seek explanation at the individual and school levels. PISA has the very challenging task of monitoring change in achievement and educational processes at the national level, with a rapidly growing body of countries participating. This task must be done to the highest standards by those who are committed to the work.

REFERENCES

Aikin, W. M. (1942) *Adventure in American Education: Volume 1. The Story of the Eight-Year Study*. New York: Harper.

Aikenhead, G. S. and Ryan, A. C. (1992) The development of a new instrument: “Views of Science-Technology-Society” (VOSTS). *Science Education,* 76, 447-491.

Anderson, L. W. and Krathwohl, D. R. et al. (eds.) (2001) *A Taxonomy for Learning, Teaching and Assessing.* New York: Longman.

Baker, D. P. and Jones, D. P. (1993) Creating gender equality: Cross-national gender stratification and mathematical performance. *Sociology of Education,* 66, 91-103.

Biggs, J. B. and Collis, K. F. (1992) *Evaluating the Quality of Learning. The SOLO Taxonomy.* New York: Academic Press.

Bloom, B. S. et al. (1956) *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain.* New York: Longmans Green.

Bloom, B. S. Krathwohl, D. R. and Masia, B. B. (1964) *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook II: Affective Domain.* New York: David McKay.

Bloom, B. S., Hastings, J. T. and Madaus, G. F. (eds.) (1971) *The Handbook of Formative and Summative Evaluation of Student Learning.* New York: McGraw Hill.

Carroll, J. P. (1963) A model of school learning. *Teachers College Record,* 64(8), 723-3.

Carroll, J. B. (1975) *The Teaching of French as a Foreign Language in Eight Countries.* Stockholm: Almqvist and Wiksell, and New York: Wiley.

Conant, J. B. (1947) *On Understanding Science.* New Haven: Yale University Press.

Comber, L. C. and Keeves, J. P. (1973) *Science Education in Nineteen Countries: An empirical study*. Vol. I. New York: Wiley.

Dahlöff, U. (1967) Relevance and fitness analysis in comparative education. In D. E. Super (ed.) *Towards a Cross-national Model of Educational Achievement in a National Economy*. New York: Teachers College, Columbia University.

Hanushek, E. A. and Kimko, D. D. (2001) Schooling, labour-force quality and the growth of nations, *American Economic Review,* 90 (3), 1184-2008.

Holmes, B. (1981) Comparative *Education: Some Considerations of Method.* London: Allen and Unwin.

Husén, T. (ed.) (1967) *International Study of Achievement in Mathematics.* (2 Volumes) Stockholm: Almqvist and Wiksell.

Keeves, J. P. (1972) *Educational Environment and Student Achievement.* Stockholm: Almqvist and Wiksell.

Keeves, J. P. (1974) The IEA Science Project. Science achievement in three countries – Australia, the Federal Republic of Germany, and the United States, In: *Implementation of Curricula in Science Education*, (pp. 158-178). Cologne, German Commission for UNESCO.

Keeves, J. P. (ed.) (1992a) *The IEA Study of Science III: Changes in Science Education and Achievement: 1970 to 1984.* Oxford: Pergamon.

Keeves, J. P. (1992b) *Learning Science in a Changing World: Cross-national Studies of Science Achievement: 1970 to 1984.* The Hague: IEA.

Keeves, J. P. and Watanabe, R. (eds.) (2003) *International Handbook of Educational Research in the Asia-Pacific Region.* Dordrecht, The Netherlands: Kluwer.

Klopfer, L. E. (1971) Evaluation of Learning in Science. In B. S. Bloom et al. (1971) *The Handbook of Formative and Summative Evaluation of Student Learning*. (pp.559-641). New York: McGraw Hill.

Kok-Aun Toh and Ngoh-Kehang Goh (2003) Reform in Science and Technology Curricula. In J. P. Keeves and R Watanabe (eds.) *International Handbook of Educational Research in the Asia-Pacific Region,* (pp. 1243-1257). Dordrecht, The Netherlands: Kluwer.

OECD (1992 et seq) *Education at a Glance: OECD Indicators.* Paris: OECD.

Popper, K. R. (1959) *The Logic of Scientific Discovery.* London: Hutchison.

Postlethwaite. T. N. and Wiley, D. E. (1992) *The IEA Study of Science II: Science Achievement in Twenty-three Countries.* Oxford: Pergamon.

Rosier, M. J. (1980) *Changes in Secondary School Mathematics in Australia. 1964-1978.* Hawthorn, Victoria: ACER.

Rosier, M. J. and Keeves, J. P. (1991) *The IEA Study of Science I; Science Education and Curricula in Twenty-three Countries.* Oxford: Pergamon.

Robitaille, D. F. and Garden, R. A. (1989) *The IEA Study of Mathematics II: Contexts and Outcomes of School Mathematics*. Oxford: Pergamon.

Shayer, M. and Adey, P. (1981) *Towards a Science of Science Teaching: Cognitive development and curriculum demand*. London: Heinemann.

Shayer, M. and Adey, P. (2002) *Learning Intelligence: Cognitive Acceleration Across the Curriculum from 5 to 15 Years.* Buckingham: Open University Press.

Spearrit, D. (2003) Comparative Education Achievement Studies. In J. P. Keeves and R Watanabe (eds.) *International Handbook of Educational Research in the Asia-Pacific Region,* (pp. 1107-1122). Dordrecht, The Netherlands: Kluwer.

Super, D. E. (ed.) (1967) *Towards a Cross-national Model of Educational Achievement in a National Economy*. New York: Teachers College, Columbia University.

Tedman, D. K. (2002) *Science, Technology and Society in Science Education.* Adelaide: Shannon Research Press.

Tuijnman, A. (2003) Educational Research for National and Regional Development. In J. P. Keeves and R Watanabe (eds.) *International Handbook of Educational Research in the Asia-Pacific Region.* (pp. 1081-1094). Dordrecht, The Netherlands: Kluwer.

Tyler, R. W. (1949) *Basic Principles of Curriculum and Instruction*. Chicago: University of Chicago Press.

Tyler, R. W. (1986) Changing concepts of educational evaluation. *International Journal of Educational Research,* 10 (1), 1-96.

Watkins, D. (2003) Student Learning: A Cross-Cultural Perspective. In J. P. Keeves and R Watanabe (eds.) *International Handbook of Educational Research in the Asia-Pacific Region,* (pp. 441-462). Dordrecht, The Netherlands: Kluwer.