doc. Ivan Rekun
Melitopol Institute of Public and Municipal Administration of the "Classic Private University"

Models for the evaluation of educational achievement

INTRODUCTION

In his keynote address to the 1989 International Conference on Information Systems, Lester Thurow suggested that Corporate America adopt a new philosophy for the 1990s, one that centers on personnel development rather than "the building of a better mousetrap." To become more competitive with the rest of the world, organizations need to view employees as assets whose value can be enhanced through education and training.

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. 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 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

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

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.

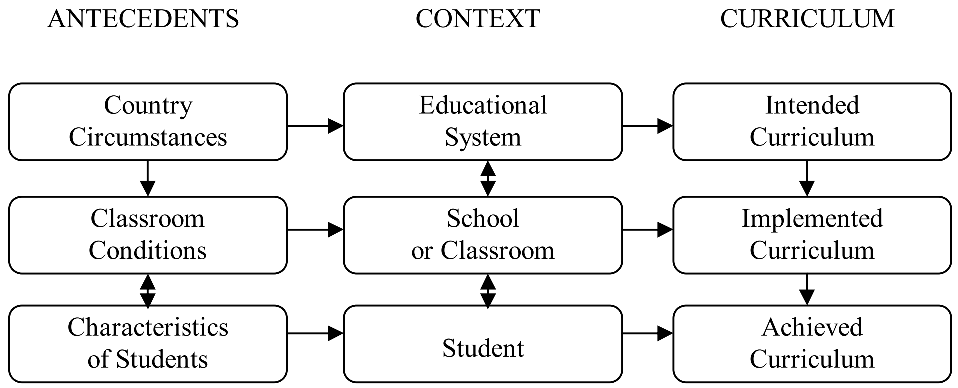

Figure 1. The context and components of the school curriculum

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 2.

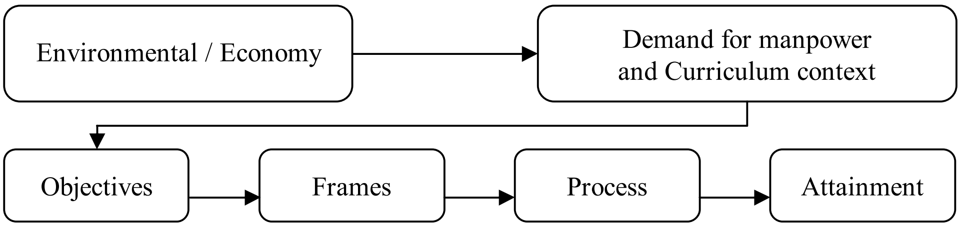

Figure 2. A cross-national model of educational achievement in a national economy

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.

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.

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), s. 1184–2008.

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

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

Rosier M.J. and Keeves J.P., 1991, *The IEA Study of Science I*; Science Education and Curricula in Twenty-three Co.

Summary

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.

Modele Oceny Osiągnięć Szkolnych

Streszczenie

W artykule omówiono zagadnienia związane z monitorowaniem procesu nauczania w zmieniającym się świecie. Autor podkreśla konieczność rozwoju międzykulturowych badań dotyczacych osiągnięć edukacyjnych, szczególnie w dziedzinie nauk ścisłych. Istotna wydaje się być również analiza podstaw teoretycznych tych badań oraz poszukiwanie najlepszych modeli do ich ewaluacji. W artykule zidentyfikowano 10 obszarów badawczych, obejmujących najbardziej krytyczne na chwilę obecną problemy dotyczące nauczania nauk ścisłych na świecie. Odwołano się w szczególności do badań w ramach projektu PISA prowadzonych przez Organizację Współpracy Gospodarczej i Rozwoju.