

Current ConcernsCurrent Concerns - The monthly journal for independent thought, ethical standards and moral responsibility.  
No 1, 2004, P.O. Box 223, CH-8044 Zurich, Phone +41-1-350 65 50, Email: editors@currentconcerns.ch

OECD: The Trojan Horse Within

Short history of the OECD and its PISA activities

Fred Naylor, Educational Consultant, UK

The Organisation of Economic Cooperation and Development (OECD) was set up in Paris in 1960. Its main aims were:

- to achieve the highest sustainable economic growth and employment and a rising standard of living in Member countries, while maintaining financial stability, and thus to contribute to the development of the world economy;
- to contribute to sound economic expansion in Member as well as non-member countries in the process of economic development; and
- to contribute to the expansion of world trade on a multi-lateral, non-discriminatory basis in accordance with international obligations.

There were 20 founding members from around the world; the number had increased to 30 by 2000. A notable feature of the OECD is its drive for expansion in membership. A significant and growing number of non-OECD countries already take part in the PISA studies.

Not surprisingly, the member countries turned their attention to educational issues during the first decade of the OECD's existence. A prime aim of a state-funded education system is to anticipate the kind of work force required by the nation's employers, and hence the kind of training that needs to be provided in schools and colleges. This, however, is only one of the aims of education. Traditionally, schools have existed to provide character training, and to ensure that morality underlies social activities.

Educational aim is equality of results

The problem of a group of nations agreeing on moral, social and indeed political matters grows as the number of nations within the group increases. Another factor that makes agreement difficult is the growing pluralism within individual member countries. This problem was a feature of the 1960s, and soon made its presence felt within the OECD - particularly in the educational context. A turning point was reached in the early 1970s. This was delineated by the Swedish professor and OECD consultant, Torsten Husen, in his ground-breaking OECD publication *Social Influences on Educational Attainment*, 1975. In it he announced that the OECD had replaced the educational aim of 'equality of opportunity' by 'equality of results'. The first he described as 'liberal'. Equality of results was depicted as 'the more "radical" or sociological (and sometimes Marxist)' concept. Since that time the political bias of the OECD has been firmly grounded. The conflict between equality of opportunity and equality of results (one cannot have both) underlies all major world issues today. Another aspect of Marxism absorbed by the OECD is Internationalism. Tony Blair's guru, Anthony Giddens, is the author of *The Third Way*, 1998. It contains a chapter entitled 'Into the Global Age', in which Giddens speaks of the OECD as 'a possible bridgehead to a wider global order if it absorbed the European Union'. He has in mind a European Union, which has taken the form of a federation and rid its members of their national identities; and whose Parliament will act as a blueprint for that which will be eventually established within the United Nations.

The bias of the OECD is not widely recognised, particularly among many national political leaders. In 1980 I drew it to the attention of Britain's Prime Minister, Margaret Thatcher. She replied, thanking me 'for letting me know about these matters and for analysing the several documents so clearly'. The first was published around the time of the 1970 General Election (after which she became Education Minister). She stated that it had never been brought to her attention. Dealing with the Husen publication, she declared: 'I thought we had got rid of that patently ridiculous theory long ago.' Unfortunately, during her period as

Prime Minister the OECD was not deflected from the course it had embarked upon. Doubly unfortunately, the patently ridiculous notion that we should strive for equality of results in education has become firmly rooted in British and some European government circles today.

The OECD PISA operations (Programme for International Student Assessment) were initiated in the late 1990s and the first results (PISA 2000) were published in 2001. These related to 28 of the OECD countries plus 4 non-OECD countries - Brazil, Latvia, Liechtenstein and the Russian Federation. Another 13 countries, including China, took part in the same survey later, and their results were published in July 2003.<sup>1</sup> The Slovak Republic and Turkey are going to be included in the PISA 2003 survey.

PISA hailed as evidence against selective system ...

There have been mixed reactions to PISA 2000, which is not surprising considering the OECD's political agenda. It has been unjustifiably hailed by left-wing supporters and Press as providing evidence supporting their attack on selective systems of education. The most striking example of this is an article (31 January 2003) in the Times Educational Supplement (TES) by a former Editor, in which she stated that the percentages of pupils reaching the 5+ GCSE 'higher grades' standard in England was higher than that reaching the corresponding standard in Germany. Not only was there no mention of GCSE examinations in the PISA report, but there is no known way of relating the standards of public examinations in the different countries. It was all a misrepresentation of the finding that Germany's position in relation to England - in mathematics - in PISA 2000 was much lower than it had been in the IEA (TIMSS)<sup>2</sup> survey carried out a year earlier. A similar example was provided by the Independent newspaper in reporting the PISA results of the 13 additional countries. It claimed that it had been shown that streaming had an adverse effect on educational achievement. An investigation, however, revealed that there was absolutely no reference to streaming in the report. The Independent had simply reported the opinion expressed by the OECD overall co-ordinator, Andreas Schleicher, who had chaired the Press Conference!

As a matter of supreme importance, in view of the opinions expressed by Schleicher, it should be noted that on page 212 of the PISA 2000 Report it was stated: 'PISA results suggest that there is no single factor that explains why some schools or some countries have better results than others. Successful performance is attributable to a constellation of factors, including school resources, school policy and practice, and classroom practice.'

... but criticized by eminent statisticians

Against these misinterpretations of the Report's findings for political reasons, PISA 2000 has come in for severe criticism from a number of eminent statisticians. Prof. Elart Von Collani of the University of Wurzburg wrote a paper entitled OECD PISA - An example of Stochastic Illiteracy?<sup>3</sup> In it the professor listed a number of errors made by PISA in designing and carrying out the exercise. He concluded that most results could not be evaluated as the necessary information about sample size, model and methods had not been made available. Almost no absolute data had been given; it had been transformed and made relative. For example, he noted that the average score in each of the literacies tested was set at 500, with a standard deviation of 100. And he drew attention to the fact that the times allowed for the regular PISA tests varied from country to country.

Background factors not accounted for

The most striking conclusion to be drawn from this kind of statistical critique was that the PISA comparisons were never of a multivariate nature. To understand the meaning of this one should compare the PISA exercise with that undertaken by the Statistical Division of the English Department of Education and Science published in 1984 ('Statistical Bulletin 13/84'). In an attempt to discover the relationship between school standards and the socio-economic and other background factors of pupils in maintained schools, the performances of pupils in the nation's 16+ public examinations were compared across the 96 local education authorities. In all, 23 background factors were investigated. By a

process known as multivariate regression analysis equations were established, which related a given level of performance to each of the background factors that were correlated with it. PISA, by contrast, whenever it did try to establish correlations, limited itself to comparing two variables at a time. In this way it failed to make allowance for background factors.

Statistical Bulletin 13/84 was notable for its finding that the percentage of pupils in grammar schools in local authorities was related to examination success in a positive way. This finding was played down in the Bulletin by a failure to describe the magnitude of the effects. With the support of the Education Minister, Sir Keith Joseph, I was able to arrange a meeting between the Department's statisticians and a group of colleagues - including Professor S.J. Prais. We reached agreement, which was published in Hansard<sup>4</sup> as a result of the asking of a Parliamentary Question. The main points agreed were that the 13/84 analysis indicated that for three out of the five categories of examination attainments studied - covering pupils across the ability range - results were significantly higher in those local authorities that had retained a selective system of schooling. This was after due allowance had been made for background factors. Furthermore, the superiority shown by a selective authority was directly proportional to the degree of selectivity it exhibited. This kind of exercise has never been repeated.

For me, Von Collani's thoughts cast a new light on the PISA procedures. There is no limit to the observations that could be made by a vast army of inspectors, dedicated to the task of observing and recording. Almost limitless sets of figures could be produced. For example, why not the number of times students switched off their attention from their teacher during a lesson? And using the questionnaire why not ask them to describe fully their smoking and drinking habits? The jungle of statistics could be dipped into and utilised by anyone who had a pre-conceived idea that needed substantiating. For example, the two most successful countries of all in PISA 2000 were Japan and Korea. The same two countries spent very much less per pupil by having the largest class sizes by a very wide margin. For example, the average secondary school class sizes of Japan and Korea were respectively 35 and 38. This was not something the OECD was interested in, so it was never remarked upon.

What they were interested in, given their commitment to equality of results, were countries in which differences between pupil performances were relatively low. They were conscious of the fact that there is a very strong argument that differences can only be minimised by handicapping the more able pupils, or by applying reverse discrimination to the less able. They were therefore always on the lookout for any instances where narrower differences were associated with higher than average performances (equality and quality, as they described it). They were careful, however, in trying to summarise their findings to leave them suitably vague. The most astounding omission from the PISA reports was any attempt to measure the innate abilities of the students. It was assumed that all differences were entirely due to the conditions under which the students worked and lived. Given the success of the current genome projects this is an assumption that will not be allowed to go unchallenged for much longer.

Another eminent statistician, who has criticised PISA 2000, is Professor S.J. Prais. He and the organisation for which he works, the National Institute of Economic and Social Research (NIESR), have been involved for some time in international comparative studies, particularly detailed comparisons between Switzerland and Britain. Their direct class observations confirmed that the previous international IEA (TIMSS) test-surveys of mathematic attainment were correct in reporting that pupils in Switzerland (and Germany) were well ahead of those in Britain. PISA studies, however, suggested that Britain might have caught up since the last TIMSS study a year previously (both countries had a PISA mean score of 529). The same considerations applied in the case of Germany, whose PISA mean score of 490 was taken by its critics as implying that it had crashed.

TIMSS survey on the same age cohort produced different findings

This led to Prof. Prais writing a paper which appeared in the Oxford Review of Education, June 2003, Cautions on OECD's Recent Educational Survey (PISA). His

criticisms were purely of a statistical, rather than political, nature. He addressed himself to the differences between the IEA (TIMSS) survey published in 1999 and the PISA report 2000. Since the IEA survey was carried out on 14 year-olds, and the PISA survey was carried out on 15 year-olds, one year later, he pointed out that the same age cohort was involved in each. The differences between the two sets of findings were therefore even more unexpected. He identified four main differences between the objectives and methods adopted in the IEA and PISA studies, which invalidated the direct comparisons that were being made in respect of their findings.

Imprecise and confusing questions

(i) Questions in the IEA survey were directed to the mastery of the school syllabus by students in the relevant school Grade, whereas PISA was deliberately directed to so-called 'every day life' problems. It was not mathematics that PISA was concerned with, but rather 'mathematical literacy'. PISA had admitted that the students were provided with written information and were expected to interpret it and draw conclusions from it. I entirely agree with Prof. Prais. I found many of the PISA questions difficult to answer, not because they were difficult from the mathematical point of view, but because they were imprecise and confusing. Let me give an example:

A seal typically comes to the surface of the water to breathe; it then dives to the bottom and rests there for 3 minutes; then slowly floats to the surface in 8 minutes; takes a breath, and repeats the cycle. Where was the seal 60 minutes after having come to the top to breathe?

So much necessary information is missing that it must put a person with mathematical ability under stress. How long does it take the seal to dive to the bottom? How long does it take the seal to fill its lungs with air? What is the difference between 'slowly floating' and 'floating'? How slowly did it float to the surface? A much more mathematical and less literacy-dominated question would have asked the student to work out what was the minimum amount of time, devoted to diving and breathing in the cycle, that would have ensured that the seal was at the bottom rather than floating upwards after 60 minutes.

Warning carefully hidden

PISA, quite properly, gave a stark warning on p.26 of PISA 2000 that 'If one country's PISA scores are higher than those of another country, it cannot automatically be inferred that the schools in the former are more effective, since learning starts well before school and occurs in a range of institutional and out-of-school settings.' This warning, however, was very carefully hidden away, and was in fact ignored by the authors of the PISA report themselves, who elsewhere stated: 'PISA seeks to compare how well different school systems prepare students for life.' It added that the results would help governments to make policy decisions. Politicians such as Tony Blair have ignored the warning, and have wrongly taken the findings as an endorsement of the success of their national educational systems.

(ii) The IEA (TIMSS) survey was based on samples of whole classes, including late entrants or repeaters. The PISA study was based strictly on a 12 month period of birth, and excluded repeaters and some late entrants. The variability within a class could not therefore be examined in the PISA survey.

(iii) England's response rate for schools was 60 per cent, compared with 95 per cent in leading European countries.

(iv) England's response rate for pupils was lower than in any other country and lower than in the previous IEA (TIMSS) survey.

No valid conclusions

Prof. Prais drew attention to the fact that the PISA returns for Hamburg and Berlin were omitted from the figures for Germany, because of poor response rates; similar omissions were not made in the case of England, despite similarly low response rates.

He also mentioned the finding in the OECD Adult Literacy Study that the French were not capable of reading their country's newspapers!

In his summing up Prof. Prais stated that it was difficult to draw valid conclusions for Britain, and planned repeats of the studies should be postponed until methodological problems had been resolved. The English taxpayer should not

continue to support PISA studies.

In an Annex to his paper he discussed the computational mystery of PISA's country-rankings. He found (to his astonishment!) that, for the results of the 31 mathematical questions that have more recently become available, Switzerland was ahead of Britain in answering correctly two-thirds of the number of questions (21/31).

To complete the story it should be said that the OECD has produced a 10 page Rejoinder to Prof. Prais' paper, and aimed to publish it in August 2003. It recognises that Prais, and it itself, may be using different criteria for judging model school systems. If it really believes its own statement that the figures should not be used to compare educational systems, then in saying that its results are sound it is not expressing any criticism of the professor. Furthermore, it admits that its PISA report had agreed (i) that the UK school response rate of 59% (prior to replacement) and 82% after replacement was a matter for concern, as was the student level response of 81%, and (ii) such low response rates brought a threat of bias that could not be formally addressed. Irrational interpretations of the Left

The most recent publication by PISA (now in association with UNESCO!) in July 2003 incorporated the results of the additional 13 countries taking part in PISA 2000. It evoked the most amazing reactions from the left-wing organisations and press in England, who fully share the OECD bias and are ready to put their own interpretations on the vague and unsound observations made by the report's authors. The Times Educational Supplement's heading was 'Class Segregation Holds Britain Back'. The Guardian led with 'Finland boasts comprehensive system where standards soar'; the Times with 'Streaming pupils "worsens results"'. John Bangs of the National Union of Teachers (NUT) claimed: 'It is the final nail in the coffin of selection in any way, shape or form'. The Association of Teachers and Lecturers (ATL) said: 'It proved state run education systems can simultaneously achieve educational quality and equity, which will be music to teachers' ears'.

Of course they had all failed to pick up the PISA 2000 Report's point on page 212 that successful performance is attributable to no single factor, but to a constellation of factors, including school resources, school policy and practice, and classroom practice.

These irrational reactions matched the left-wing reaction to the lowly positions of Germany and Switzerland in PISA 2000 (published in 2001). They were wrongfully taken as an indication that the tripartite educational system, for which these countries are famous, was less successful than the comprehensive system. The German teachers of mathematics and science were livid at these misinterpretations, and revolted against the type of questions used. No attempt has been made by the PISA researchers to compare the performances of the comprehensive and selective systems in any of the countries in which they exist side by side. The result is that Germany is reacting by carrying out research on the performances of the individual Lander by supplementing the PISA 'literacy' tests with academic tests more suited to the school curriculum. The Swiss apparently have not reacted as sharply as Germany.

Selective system superior

In England the superiority of the selective system is supported by evidence especially suited for the purpose. Not only do we have the results of the impeccable 13/84 statistical analysis The statistics published annually for success in the GCSE have related only to the top 50 per cent of the population, but recent figures extracted from the Government by the Conservative Party in respect of higher ability pupils demonstrate the great superiority of schools in the selective system. The combined results of grammar and secondary modern schools are nearly double those in the comprehensives. For example, in wholly selective areas 15.1% of pupils achieved five or more GCSE grades A\*-A in 2002. The corresponding figure in wholly comprehensive areas was 8.6%.

The idea that the relative strengths of the selective and comprehensive systems can be deduced by ignoring the evidence collected on this subject in the countries in which the two systems exist side by side, and instead using the mass of unrelated evidence collected from countries world wide, is bizarre. It

is equally bizarre to expect such politically diverse countries as the USA, the Russian Federation and China to be making their policy decisions on the statistics gathered by PISA in the context of its own distinctive educational model. It is noteworthy that, in criticising Prof. Prais' criteria for citing Switzerland as a model educational system within the European context, the PISA Rejoinder defines the criteria on which it bases its own ideas of a model system. These are 'the level and distribution of educational outcomes.' Our present Secretary of State for Education has made a plea that we should consider the performances of our different systems by the scientific evidence alone, uninfluenced by ideological considerations. This is an excellent recipe, and it is clear that Britain and its European counterparts should join forces to demonstrate scientifically that selective systems of education, based on equality of opportunity, produce better academic results than the comprehensive schools, based on the patently ridiculous notion that equality of results is the correct educational target. This is becoming a more urgent task day by day.

1 Literacy Skills for the World of Tomorrow, July 2003. OECD/UNESCO.

2 IEA - International Association for the Evaluation of Educational Attainment  
TIMSS - Third International Maths and Science Study (1996).

3 Economic Quality Control, 2001

4 Hansard, 15 May 1987, Columns 458, 465 and 466.

OECD ADDENDUM

(5 January 2004)

On 3 December 2003 I attended a London conference on Statistical Perspectives on PISA, held at The Royal Statistical Society. The purpose of the conference was to look critically at PISA 2000 and other related international studies. Prominent amongst the critics was Harvey Goldstein, of the Institute of Education at the University of London, who was one of the initiators of PISA. He raised methodological concerns about the analysis and interpretation of the results. He pointed to features that raised doubts about the adequacy of the data.

In particular, he spoke of the difficulty of comparing the results of culturally and linguistically different cultures, and complained that PISA relied on THE ITEM RESPONSE MODEL (IRM) to decide whether items fitted their unidimensional models. This ensured that the items chosen conformed to a given scale, and this resulted in 'dodgy items'. He also criticised PISA's na-ve approach to causation, stressing that it was essential to introduce a longitudinal component. In other words, it was necessary to examine the performances of individual countries over a number of years before any judgments could be made about their relative successes.

Other speakers were equally suspicious of PISA's findings. For example, Prof. Margaret Brown of King's College London, dealing with mathematics performances, stated that: 'The use, and sometimes even the collection, of these comparative data are often ideologically driven, either by governments trying to justify greater central control or by governments, groups or individuals arguing for changes in teaching methods or curriculum. In some of these cases, the references to the data have been selective and/or misleading.'

The speaker representing the Department for Education and Skills (DfES) defined what he termed the main weakness of PISA studies as 'the failure to measure innate qualities'. He thereby confirmed what I had earlier described as 'the most astounding omission from the PISA reports'.

One of the Discussion Sessions, led by Professor Peter Tymms, University of Durham was entitled: 'The difficulty of interpreting international achievement data in the absence of prior cognitive measures'. It summed up the general feeling of the inadequacy of the PISA studies.

Article published on 01-26-2004