SCOTTISH MATHEMATICAL COUNCIL DISCUSSION PAPER

The 2012 PISA Findings and the Implications for Scotland

Dr Chris Pritchard, SMC Chair

Background

PISA (Programme for International Student Assessment) and TIMSS (Trends in International Mathematics and Science Study) are the main studies by which the performance of pupils in mathematics is compared across nations. Scotland participated in TIMSS in 1995, 1999, 2003 and 2007 but declined the opportunity in 2011, and there are no plans to participate further. (Additionally, SSLN surveys performance in language and numeracy across Scotland in alternate years.) PISA is an assessment of the skills of 15 year-olds (typically S4 pupils in Scotland) carried out under the auspices of the Organisation for Economic Co-operation and Development (OECD). The programme runs every three years across all OECD members and a variety of partner countries (65 jurisdictions in 2012). Scotland has participated since the first wave of testing in 2000 when our scores were subsumed in the UK figures. The three areas under study (mathematics, science and reading) are each given prominence every third study, i.e. every nine years. This fifth study had a focus on mathematics.

Over half a million pupils took part in the 2012 study on which the report was published in 2013. Summaries of the performance of pupils in PISA 2012 across the world, in the UK, and in Scotland specifically can be found at:

www.oecd.org/pisa/keyfindings/pisa-2012-results.htm
www.oecd.org/unitedkingdom/PISA-2012-results-UK.pdf
www.scotland.gov.uk/Topics/Statistics/Browse/School-Education/PISA

The full report runs to 4 volumes (560 pp on performance, 332 pp on equity issues; 521 pp on pupil attitudes etc; 542 pp on school resourcing and organization). It is available via:

www.oecd.org/pisa/

Further information about PISA in Scotland may be sought from Mal Cooke (0131 244 1689, school.stats@scotland.gsi.gov.uk).

Assessment Items

As in other countries, the Scottish pupils (almost 3000 of them) tackled a 2-hour paper-based test with a mixture of open-ended and multiple-choice questions that were organised in groups based on a passage setting out a real-life situation. The layout is very similar to that used in the 5-14 assessments in Scotland. Within the themes used, individual questions are graded into six levels of difficulty. They range from routine questions (Level 1) to demanding questions requiring advanced problem solving skills and sophisticated reasoning (Level 6). Critically, rote learning (drill) is wholly inappropriate for tackling many of them and this gives the lie to critics who believe that it is simply the hours the pupils in Pacific Rim countries put in that accounts for their outstanding performance. Here are some of the questions; many more, together with detailed statements of the levels, can be found via:

www.scotland.gov.uk/Topics/Statistics/Browse/School-Education/PISA/mathquestions

Exchange Rate (Q1 at Level 1, Q2 at Level 4)

Mei-Ling from Singapore was preparing to go to South Africa for 3 months as an exchange student. She needed to change some Singapore dollars (SGD) into South African rand (ZAR).

1. Mei-Ling found out that the exchange rate between Singapore dollars and South African rand was:
1 SGD = 4.2 ZAR

Mei-Ling changed 3000 Singapore dollars into South African rand at this exchange rate. How much money in South African rand did Mei-Ling get?

2. During these 3 months the exchange rate had changed from 4.2 to 4.0 ZAR per SGD.

Was it in Mei-Ling's favour that the exchange rate now was 4.0 ZAR instead of 4.2 ZAR, when she changed her South African rand back to Singapore dollars? Give an explanation to support your answer.

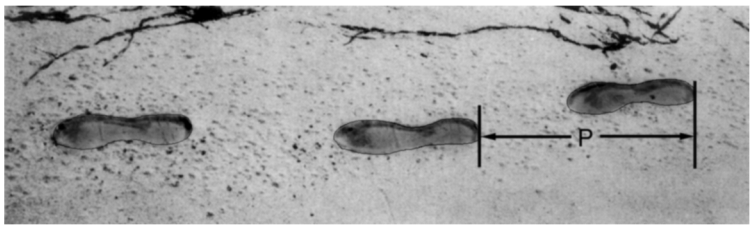
Earthquake (Level not stated)

A documentary was broadcast about earthquakes and how often earthquakes occur. It included a discussion about the predictability of earthquakes. A geologist stated: "In the next twenty years, the chance that an earthquake will occur in Zed City is two out of three".

Which of the following best reflects the meaning *of the geologist's statement*?

A 2/3 x 20 = 13.3, so between 13 and 14 years from now there will be an earthquake in Zed City.
B 2/3 is more than 1/2, so you can be sure there will be an earthquake in Zed City at some time during the next 20 years.
C The likelihood that there will be an earthquake in Zed City at some time during the next 20 years is higher than the likelihood of no earthquake.
D You cannot tell what will happen, because nobody can be sure when an earthquake will occur.

Walking (Levels not stated)



The picture shows the footprints of a man walking. The pace length *P* is the distance between the rear of two consecutive footprints.

For men, the formula, n / P =140, gives an approximate relationship between *n* and *P* where,

n = number of steps per minute,

P = pace length in metres.

1. If the formula applies to Heiko's walking and Heiko takes 70 steps per minute, what is Heiko's pace length? Show your work.

2. Bernard knows his pacelength is 0.80 metres. The formula applies to Bernard's walking. Calculate Bernard's walking speed in metres per minute and in kilometres per hour. Show your working out.

Scotland's Performance

The performance of British pupils was middling, much as in previous studies. Scottish and English pupils scored similarly, with those in Northern Ireland and especially Wales doing worse.

1. Mean: Scotland's mean scores over the four most recent studies are:

	2003	2006	2009	2012
Scotland	524	506	499	498
OECD	500	498	496	494

Across participating countries, mean scores in PISA have fallen over time but Scotland's mean has been falling by 1 point per 3-year cycle more than the OECD average. That said, the rate of

decline has now fallen to near zero and this is taken by the Scottish Government to be reassuring. (**Question**: Why did Scotland's performance nosedive between 2003 and 2006?)

2. **Levels Achieved**: The proportion of pupils across all jurisdictions able to tackle questions at each level are:

Highest Level	6	5	4	3	2	1	Below 1
Percentage	3.3	9.3	18.2	23.7	22.5	15.0	8.0

Scotland had higher proportions in the central categories, fewer at the extremes. This may or may not have contributed to …

3. **Variation**: A characteristic of Scotland's performance in 2003 and 2006 was the unusually small interquartile range. In 2009, the standard deviation was used as the measure of dispersion and again the figure was low but it fell even further in 2012 to the fifth lowest amongst the participating countries.

4. **Gender**: The gender differential in favour of boys was 7 points in 2003, 11 points in 2006, 14 points in 2009, 15 points in 2012. This last figure is largely consistent with the gender gap elsewhere. However, the gender gap in favour of boys was significantly wider among high achievers than in other countries. **Questions**: Could able girls do even better? Should a higher proportion of girls be performing with the high-attainers?

5. **Disadvantage**: The performance gap between disadvantaged and less disadvantaged Scottish pupils in mathematics appears to be in decline (37 points in 2012, 45 points in 2009). This still amounts to 1 year of learning, e.g. a socially disadvantaged 16-year old operating at the level of 15-year old with no such deprivation.

6. **Content categories**: PISA items fall into one of four mathematical content categories: Quantity, Uncertainty and Data, Change and Relationships, Space and Shape. It was in the last of these categories that, despite outperforming other UK jurisdictions, our score of 482 fell below the OECD average of 490. This category is defined broadly:

> Space and shape encompasses a wide range of phenomena that are encountered everywhere: patterns, properties of objects, positions and orientations, representations of objects, decoding and encoding of visual information, navigation, and dynamic interaction with real shapes and their representations. Geometry is essential to space and shape, but the category extends beyond traditional geometry in content, meaning and method, drawing on elements of other mathematical areas, such as spatial visualisation, measurement and algebra. Mathematical literacy in space and shape involves understanding perspective, creating and reading maps, transforming shapes with and without technology, interpreting views of three dimensional scenes from various perspectives, and constructing representations of shapes.

Other Issues

1. Teachers' Pay and Performance in PISA

Amongst the developed countries, there appears to be a positive correlation between teachers' pay and performance in PISA, with countries scoring 10 points higher typically receiving 30% above Scotland's teachers. (See pages 26-27 of *PISA 2012 Results in Focus* at www.oecd.org/pisa/keyfindings/pisa-2012-results-overview.pdf.) This association is not causal and hence there is no direct implication that raising pay will help boost performance. On the other hand, it is conceivable that teacher pay, status and motivation are causally associated. Should we be concerned about falling levels of motivation amongst Scottish mathematics teachers as they pass through several years of pay erosion and curricular innovation, and could there be an adverse effect on performance in PISA 2015?

2. Social Disadvantage, School Resourcing and Performance in PISA

The 2012 survey also revealed an association between a country's performance in PISA and the extent to which it was able through its policies on resourcing to redress the social disadvantage

experienced in some areas and by some communities. (See pages 27-28 of *PISA 2012 Results in Focus*.) Finland appears to be a leader in this regard (schools function fairly uniformly across the country with the overwhelming majority of youngsters receiving high-quality and highly-resourced education), whilst USA do particularly poorly. Without appearing complacent, it is fair to say that Scotland does rather well in this regard, as mentioned earlier.

Criticisms of PISA

David Leat (Newcastle University) notes that 'most politicians take the results very seriously, making all manner of interpretation, perhaps because the popular press makes the results into simple headlines', though 'most academics, analysts and some heavyweight newspapers are sceptical and critical about the results and simple interpretations, although they do not completely dismiss the data.' The Guardian correspondent, Peter Wilby, wrote an article on 1 December 2013 entitled: *'*Don't let dubious league tables dictate how we educate our children*'.* Here he argued that

> there are ample reasons not only to question whether average scores from written tests can adequately assess the quality of school systems across the planet, but also to argue that international testing regimes pose a threat to national sovereignty and cultural diversity.

Australian educationalists, led by Jennifer Buckingham, are also sceptical. In December 2012, she wrote 'Keeping PISA in perspective; Why Australian education policy should not be driven by international test results'. One point she raises is that PISA questions are about mathematics in a real-world context and do not test what is met in the classroom while TIMSS questions more closely reflect what is learnt in lessons. The former assesses the skills a person has gained from all quarters while the latter assesses what has been learnt at school.

Questions

1. Will the advent of *Curriculum for Excellence*, with its greater emphasis on mathematics in context, do something to improve Scotland's PISA results subsequently?
2. Is there significant overlap between the kinds of questions in PISA and those in the SMC Mathematical Challenge? Can we justifiably promote MC on the grounds that, in part, it develops skills tested by PISA?

There are three subheadings under which criticism of PISA might be arranged:

- **Gaming**: There is evidence that some countries 'game the system' in ways which Scotland should not emulate. For example, Poland has redrafted its curriculum into a style conducive to doing well in PISA studies. (This may be a point missed by Alasdair Allan, Minister for Learning, who has identified Poland as a country to watch.) The OECD is aware that Germany, Switzerland and Spain are known to have gamed the system. China has taken an alternative approach to gaming, publishing the PISA scores from just its highest-performing jurisdiction, Shanghai, where educational provision is known to be far higher than elsewhere in the country.
- **Cultural Differences**: There appears to be a correlation between reading skills and performance on the PISA items and that this, for example, depresses scores in English speaking countries whilst bolstering scores in Finland and Korea where the languages are phonetic. Jennifer Buckingham goes so far as to claim that

  > The PISA maths test is more highly correlated with the PISA reading test than with the maths test in the Trends in International Maths and Science Study (TIMSS), suggesting that reading ability is a better predictor of PISA maths scores than maths knowledge.

  (The data from TIMSS are influenced to a lesser extent in this regard.) Part of the issue in Finland is that society is very homogeneous, with virtually no immigration. Meanwhile, Finnish university lecturers complain about the sharply declining skills of their intake year-on-year, a situation which has arisen because PISA skills are practised at the expense of those skills needed for later study. (Their TIMSS scores are not strong.) Finland's neighbour, Sweden, does have immigration and falling results. The poor performance here might be associated with the policy of promoting the opening of free schools which is now being hurriedly abondoned. Pacific Rim countries rely for high levels of performance on after-school and weekend cramming classes. There is a climate of diligence and obedience, but a lack of creative criticism. Western society operates differently. Worryingly, Indonesia is at the bottom of the league tables, yet it has the highest percentage of children who say that they are happy at school; South Korea has almost the reverse pattern, together with the highest rate of teenage suicides.

- **Flaws in the Statistics**: The statistical model on which PISA rests is mathematically flawed and the rankings arising from the model are worthless as a result (Hugh Morrison, Queen's University Belfast). There are also problems with the statistics included (some 16-year olds repeating a year are included in some jurisdictions) or the substitution of estimated values. In Denmark they have argued that when the balance between the topics covered is adjusted, the result are changed radically. In other words, the outcomes are sensitive to the issue of topic content. Finally, in trying to reduce the effects of bias, there is an overly strong emphasis on commenting on common aspects, effectively throwing away information on interesting areas of difference (Harvey Goldstein, Bristol University).

Looking to Other Countries

With the publication of each set of international rankings, there is a rush to see which countries around the world we might learn from. Much has been made of Finland's consistently good showing but below, and elsewhere in this document, there is some evidence that the picture in Finland is not quite as rosy as first thought. Yet politicians might wish us to perform at least a dozen points higher than currently but not at levels peculiar to the Pacific Rim where pupil experiences are so very different:

Germany	514	(Finland	519)
Belgium	515	Estonia	521
Canada	518	Netherlands	523
Poland	518		

Good countries for Scotland to emulate are Belgium and the Netherlands. In comparison to Scotland they are a little larger (at twice and three times our population), western democracies which are geographically and culturally not too distant. Paul Andrews has previously argued that the UK should look to Flanders (specifically Brabant) as a model of successful learning and teaching in mathematics. See: Andrews, P. (2013). What does PISA performance tell us about mathematics teaching quality? Case studies from Finland and Flanders. In H.-D. Meyer & A. Benavot (Eds*.), Who succeeds at PISA and why? The role of international benchmarking in the emerging global education governance system. Institutional and policy perspectives* (pp. 99-114). Oxford: Symposium. Part of the abstract reads:

> Flanders, whose PISA achievements have been masked by those of Belgium as a whole, has performed as well as Finland with respect to mathematical literacy and, on the three TIMSS in which it has participated, it has been the most successful European system at grade 8. … Flemish didactics may have greater explanatory potential for both PISA and TIMSS success. Such findings suggest that performance on international tests of achievement may be unrelated to didactical quality as other, typically hidden, cultural factors intercede.

A more detailed look at the figures for Belgium reveal that though the overall mean is 515 (boys 520, girls 509), the figures for the Flemish community alone are 531 (537, 525). And in the 'Space and Shape' category, Belgium scored 509 in this category, Netherlands 507, both well above the average.

Lessons from Shanghai

In England, mathematics teachers are being urged to learn from high-performing Shanghai. The Department for Education has issued a *Report on research into maths and science teaching in the Shanghai region* suggesting how Shanghai's approaches could be implemented in schools in England. (www.education.gov.uk/nationalcollege/docinfo?id=178400&filename=report-on-research-into-maths-and-science-teaching-in-the-shanghai-region.pdf).

The performance of Shanghai's students in PISA was outstanding. If we exclude special schools, there are over 700 schools in the city and it was from these schools that a stratified sample was taken. Rachael Horsman (a presenter at SMC Conference 2014, though not addressing this subject) visited Shanghai in early January, and learnt about how very different learning and teaching mathematics is from the UK. She noted that:

1. Of all school subjects, Mathematics has the greatest priority and highest status, followed by English Language and then Chinese.

2. Each mathematics teacher has just one class which they see twice a day, first period and last. They therefore teach just 10 periods per week. Every day, homework is set during the last period and it is handed in at the end of Period 1 the next day, only to be corrected by the end of the day. This is where students get their practice. Teachers spend the rest of the intervening time in collegiate lesson planning (also found in Japanese schools). The teacher's school day typically finishes about 6 pm but may be as late at 9 pm once a week.
3. Teaching is pitched at the top of the class, not the middle, so that differentiation (not seen in the lesson but organised after the lesson) is downwards only. Out of respect for the teachers they are not questioned by students during the lesson even when understanding falters. Instead, the teacher is approached at the end of the period for advice and goes away after school (often to additional provision).
4. Much of the lesson is given over to tackling questions, presented in sequence. No two questions are alike; every question has a new twist, effectively making it what we would term a 'lever question'.

In Scotland, we tend to dismiss out of hand what is happening in the Pacific Rim because of the very significant cultural differences. However, we should continue to encourage the raising of the status of Mathematics, arguing for high levels of resourcing in terms of time, staffing and finance. Adjusting the school day in any but a minor way appears unlikely but regimes in which short homeworks are set regularly could be implemented. Likewise, pitching lessons towards the top of the class is not feasible but pitching them but at a higher level than currently is possible, as is reducing the number of routine questions in favour of greater problem solving.

Final Thoughts

1. Different high-scoring jurisdictions organize and teach mathematics in totally different ways. The Pacific Rim countries often have very large classes supplemented by cramming sessions after school or at weekends. This is stressful and boring for children. In Finland, they have small classes. But all profess to value education in a way which used to be professed with enormous energy in Scotland. It appears that national culture matters more than what we actually do. A return to the public having significant faith in mathematics teachers is needed. In fact, we need a teaching force with high-level skills and post-graduate qualifications, strongly supported and respected by headteachers, parents and politicians. We need to reduce the focus on management qualifications and develop middle managers with real expertise in learning and teaching mathematics.

2. There is a strong argument to the effect that in focussing on the rankings we are misusing the PISA data. What they actually allow us to do is to draw conclusions about the impact that different factors have on outcomes in education systems, helping us to identify links between school structures, teacher qualifications, socio-economic background and non-native speakers. (David Raffe of Edinburgh University has argued that the most important role the 2012 statistics will perform is as a baseline for assessing the effects of *Curriculum for Excellence*. If this is so, a shock could be in store because the embedding of mathematics in project-style work is reducing the time spent of practising routines. On the other hand, if the SQA were to take the opportunity to incorporate more problem solving in their papers, returning to something nearer the style of Credit rather than Intermediate 2, there could also be an effect in a positive direction.) The data also help to track what is happening in a particular country over time. If there is an inherent overestimation of how well a country is doing prior to a study (Germany's awful performance in the 2000 study revealed such complacency) or a significant drop in performance (Sweden dropped from 500 in 2009 to 478 in 2012), then education policy should be re-examined and revamped.

3. TESS reports that 'a Pisa-based test for individual schools will be available from next year in the UK, the US and Spain'. This was written in December 2013, so 'next year' means 2014. It seems likely that schools that can afford to purchase the package will do so and fight it out for top spot. Like Poland, Wales is attempting to revamp learning and teaching in a bespoke way in order to close the gap in its PISA scores relative to the rest of the UK. Perhaps we should be wary of the same approach being used here in Scotland.