
Chair of the UK Statistics Authority, Andrew Dilnot CBE

Rt. Hon. David Miliband MP
House of Commons
LONDON
SW1A 0AA

3 October 2012

Dear David

PROGRAMME FOR INTERNATIONAL STUDENT ASSESSMENT (PISA) STATISTICS

Thank you for your letter of 3 September. I apologise for the delay in responding. You asked me about comments made by Sir Michael Wilshaw, HM Chief Inspector of Education, in a recent BBC television interview, drawing on educational attainment data in the Programme for International Student Assessment (PISA) study. I note that PISA data for England and the United Kingdom have been reported in various sources, including the National Foundation for Educational Research (NFER), the Organisation for Economic Cooperation and Development (OECD), and the Department for Education itself. The OECD's 2009 report for the UK included the following important caveat:

"Trend comparisons, which are a feature of the PISA 2009 reporting are not reported here because for the United Kingdom it is only possible to compare 2006 and 2009 data. As the PISA 2000 and PISA 2003 samples for the United Kingdom did not meet the PISA response-rate standards, no trend comparisons are possible for these years."ⁱ

The NFER's PISA study report for England provided advice on the interpretation of country rankings.ⁱⁱ I have also noted a review published by the Institute for Education which concluded that there were problems with identifying change over time using PISA data for England, that conclusions should not be based on this resource alone, and that other evidence (including the Trends in International Mathematics and Science Study (TIMSS) to which you refer) contradicts the findings of the PISA study and therefore it may be difficult to treat an apparent decline in secondary school pupils' performance as "a statistically robust result".ⁱⁱⁱ

Given these various published caveats and advice, my own view is that those commenting on data in this area should take particular care to avoid making comparisons which could be interpreted as statistically problematic, or otherwise might be misunderstood by the audience.

I was concerned to review the Department for Education's press release of 7 December 2010 in which headline results for England from the PISA study, alongside relative international rankings, were not accompanied by detailed advice or caveats to help the reader in making comparisons over time, nor were the statistical implications of an increase in the number of reporting countries in later PISA studies noted.^{iv} While I understand that

some users of these data would like to make comparisons between the first PISA study in 2000 and the most recent in 2009, the weaknesses relating to the response-rate standard in earlier studies should not be ignored. The validity of comparisons of national rankings as a result of an increase in the number of countries covered by the PISA study, and the degrees of uncertainty in country scores attributed to sampling and measurement error are also important in this regard. I have also noted the Statistical Release published by the Department on the same date, 7 December 2010, which compares 2009 and 2006 data, but does not attempt comparisons with the earlier years, 2003 and 2000.^v

Guidance for officials circulated to departments by the former Cabinet Secretary in 2009 recommended that, when preparing publications containing statistics, the advice of statistical professionals should be sought at the earliest opportunity, and information relating to their quality, reliability and usability should be provided.^{vi}

While the PISA results themselves are not official statistics because they are a result of an international research study by organisations not covered by the scope of the *Statistics and Registration Service Act*, the Department for Education is a producer of official statistics. Given the 2009 guidance, it would have been helpful had the Department's December 2010 press release attempted to meet standards similar to those required of departmental statistical publications.

The principles of good statistical communication and statistical literacy emphasise the importance of recognising statistical uncertainty and quality weaknesses in the commentary and advice provided to the user. These uncertainties and weaknesses are not just a technical footnote; they are themselves an important part of the evidence, and affect interpretation and meaning. League tables and the presentation of international rankings can be statistically problematic, and require clear and careful commentary alongside them. I will find a suitable opportunity to explore these matters further with the Department.

I am copying this letter to the Chief Inspector, Sir Michael Wilshaw, to Jil Matheson, the National Statistician, and to the Heads of Profession for Statistics at the Department for Education and Ofsted.

With best wishes.

Yours sincerely



Andrew Dilnot CBE

ⁱ <http://www.oecd.org/pisa/46624007.pdf>

ⁱⁱ http://www.nfer.ac.uk/nfer/publications/NPDZ01/NPDZ01_home.cfm?publicationID=549&title=PISA%202009:%20Achievement%20of%2015-year-olds%20in%20England

ⁱⁱⁱ <http://www.ioe.ac.uk/60021.html>

^{iv} <http://www.education.gov.uk/inthenews/inthenews/a0070008/secretary-of-state-comments-on-pisa-study-of-school-systems>

^v <http://media.education.gov.uk/assets/files/pdf/osr292010pdf.pdf>

^{vi} <http://www.parliament.uk/documents/upload/letter-brennan-090227.pdf>