Dubious clues?

Comments by Jane Galbraith on 'Statistical clues to social injustice' by Danny Dorling (Radical Statistics 102)

By misrepresenting the research he criticises, Danny Dorling (2010) weakens his argument that the way statistical information is packaged and presented can provide unjustified support for the unequal distribution of resources.

Danny, p 43, emphasises the importance of the author giving full references and of the reader checking them. Unfortunately to read every reference is too time-consuming for this reader and, in the case of the Pearson 1895 paper, too difficult! But I have read parts of the PISA report (OECD, 2007) and technical report (OECD, 2009) and I have struggled with the terminology and mathematics in Pearson (1895) and I judge that Danny has misrepresented them. I have also re-read the discussion of John Snow's evidence given in Tufte (1997). Unlike Danny, Tufte endorses Snow's arguments.

With respect to the PISA study, Danny gives his own one word summaries of the seven categories of competence in science in Figure 1, p 46, with longer versions given underneath. These can be compared with the descriptions for Levels 1 to 6 given in the PISA report and reproduced, slightly annotated, in Danny's Appendix, pp 60-62. Danny has added his own description for Level 0, that is, for 15 year olds who fail to reach Level 1. I find Danny has seriously distorted the PISA descriptions, applied them to the students rather than to their competencies, and then suggested such ridiculous implications as that 61% of Dutch children are dullards. I have not read all the PISA report but in what I have read there is no suggestion that the 15 year olds in the lower categories are to be regarded as stupid or incapable.

For example, the PISA report describes Level 3 students as being able to "select facts and knowledge to explain phenomena and apply simple models or inquiry strategies" and continues, "Students at this level can interpret and use scientific concepts...". Danny labels them "Simple". How could he?

Again PISA says, "At Level 2 students have adequate scientific knowledge to provide possible explanations...." Why did Danny label them "Barely"?

Danny labels Level 6 as 'Advanced' and, under Figure 1, writes, "only 'Advanced' pupils are found to be capable, it is said, of the kind of thinking that might include 'critical insight'". This is incorrect, the PISA report described Level 5 students as being able to "bring critical insights to situations".

So, please compare for yourself what the authors of the PISA report wrote (given in Danny's appendix) with Danny's interpretations in and underneath Figure 1 and in the accompanying text. Danny expects you to agree with his interpretations but I do not!

I do not know whether Danny considers all educational testing inappropriate or whether he just objects to the PISA study methodology. Unfortunately he also misrepresents this methodology which he describes thus: "They did what they thought they should do – when they saw that the raw data that they got was a mess – they 'remodelled it', they made it fit to a normal distribution."

I have skimmed through the Technical Report (OECD, 2009) and can confirm that the data collection, manipulation and data analysis are extraordinarily complicated, but that all the methodology appears to be standard. In particular the use of latent variable modelling with a normal prior is widely used in educational testing. This typically yields scores that have a distribution close to but not identical to a normal distribution. I am not sure whether further scaling has been used to give a better approximation to the normal distribution.

The scale used for a set of scores based on a latent variable analysis can be chosen arbitrarily and should not affect the interpretation of the results (see, for example Bartholomew, 2004, chapter 8). However, Danny may be correct that by choosing an approximately normal distribution the PISA report reinforces conventional wisdom that the bulk of the population have middling ability and only a few have high or low ability. But, that does not imply that those with more ability deserve better any more than it would have if the scores had been scaled to have a uniform distribution (or been presented with the same number in each category). Furthermore, I cannot see why Danny thought the distribution of competence, as measured in the PISA study, has been or could be used to justify injustice. Why should unequal ability justify unequal treatment?

This brings me to Danny's misinterpretation of Pearson's example of the distribution of 'paupers'. Danny's Figure 4, p 51, is based on a

figure in Pearson (1895). Pearson used this example to illustrate that a point binomial polygon (not a binomial distribution as Danny describes it) fitted the data better than a normal distribution. Pearson argued that the distribution of the data, being skew, was *not normal*.

Pearson may have believed in the heritability of intelligence but nothing in Pearson's article supports Danny's suggestion that, in this example, a "bell curve" was drawn "to try to further its creator's arguments about the inheritability of intelligence".

Pearson (1895) states that the data came from Appendix 1 of Charles Booth's 'Aged Poor; Condition'. I have not pursued this reference but unless he has altered the data its approximation to a 'bell curve' cannot be blamed on Pearson. In any case I cannot see the relevance of the shape of the distribution of paupers, be it normal and bell shaped or skew, to the question of whether intelligence is inherited (and, if so, how much is genetic and how much environmental).

I was puzzled to read Danny's assertion that John Snow's evidence appears to be fabricated! Danny does not reference, either Snow (1855) or any work disputing the conclusion that the Broad Street pump was a major factor in the spread of the disease. I have only read the extracts from Snow (1885) given in Tufte (1997) but it is clear that Snow was aware that the epidemic was in decline before the pump handle had been removed. The epidemic started on 31st August 1854 with large numbers of deaths in the first four days of September after which they declined steadily. The pump handle was removed on September 7th which may or may not have prevented a second outbreak. Tufte explains that Snow's essential evidence, given in the famous map, was supported by the detailed written account of further evidence collected by Snow in determining what pump had been used by those affected and also that the nearby brewery (no cases) and work house (few cases) did not use water from the Broad Street pump.

I have not investigated Danny's comments on school leaving age and university entrance, sweet peas, skull size, penis length, worldwide income distribution, growth in GPD, household's ability to get by on their income, the share of income received by the richest 1%, or US mortgage debt, but my confidence in what he says and the conclusions he draws from these examples is undermined. I regret that he has not made a better case against those who claim to justify treating people unequally.

Danny clearly points out that some intentional or unintentional subjectivity will be involved in selecting and presenting information. But his misrepresentation of the work of others cannot be justified by the argument that it is what he understood them to mean or by the argument that it is in a good cause, even if that cause is the fight against injustice.

References

Bartholomew, DJ (2004), 'Measuring intelligence; facts and fallacies'.

Dorling, Danny (2010), 'Statistical clues to social injustice', Radical Statistics 102, pp 43-62.

OECD (2007). The Programme for International Student Assessment. Paris, Organisation for Economic Co-Operation and Development's latest PISA study of learning skills among 15-year-olds.

OECD (2009). PISA 2006 Technical Report. Paris, Organisation for Economic Co-Operation and Development's technical report on the latest PISA study of learning skills among 15-year-olds.

Pearson, K (1895) 'Contributions to the mathematical theory of evolution – II. Skew variation in homogeneous material', Philosophical transactions of the Royal Society of London, Series A, Mathematical, vol 186, pp 343-414.

Snow, John (1885) 'On the mode of communication of cholera', quoted in Tufte (1997).

Tufte, ER (1997) 'Visual explanantions'.