

Rasch Measurement: a response to Goldstein

Panayiotis Panayides^{}, Colin Robinson[#] & Peter Tymms^{*}*

^{}Durham University, UK*

[#]Freelance

We are pleased to have been given this opportunity to respond to Goldstein (2010) now that the BERJ has re-published his comments on Payanides et al (2010). A series of very clear and concise rebuttals have already been made by Linacre and Fisher (2012). We will therefore restrict our reply to four key issues.

Unidimensionality

Goldstein argues that the unidimensionality assumption means that, "while items may differ in ability [presumably meaning that they measure different abilities] there is only one ability that characterises an individual that determines that individual's response to each item" (pp. 3-4) implying that data are never unidimensional. He gives an example where a 2-dimensional set of algebra and geometry items appeared to conform to the Rasch model.

All data are multidimensional to some extent. Many psychometricians including Hambleton et al., (1991), Masters & Keeves (1999) Smith Jr. (2004) and Wright and Linacre (1989) have made it clear that unidimensionality does not implicitly mean only one factor or dimension but rather the presence of a dominant dimension and possibly of minor dimensions which do not affect the dominant one. The unidimensionality of the model is merely a reflection of the assumed unidimensionality of the majority of assessments we use. In most tests and examinations we assume that a higher mark indicates that the candidate has more of whatever it is we are trying to measure – i.e. the latent trait.

The Rasch model constructs a unidimensional framework and "multidimensionality only becomes a real concern when there are response patterns in the data indicating that the data represent two or more dimensions so disparate that it is no longer clear what latent dimension the Rasch dimension operationalizes" (Linacre, 1998, pp. 5-6).

Principal Components Analysis of the standardised Rasch residuals can show us whether these two dimensions are so dissimilar that they can be considered as different dimensions, or whether they can be considered as one. One can sensibly claim that algebraic and geometrical abilities are sub-components of mathematical ability.

Distribution assumptions

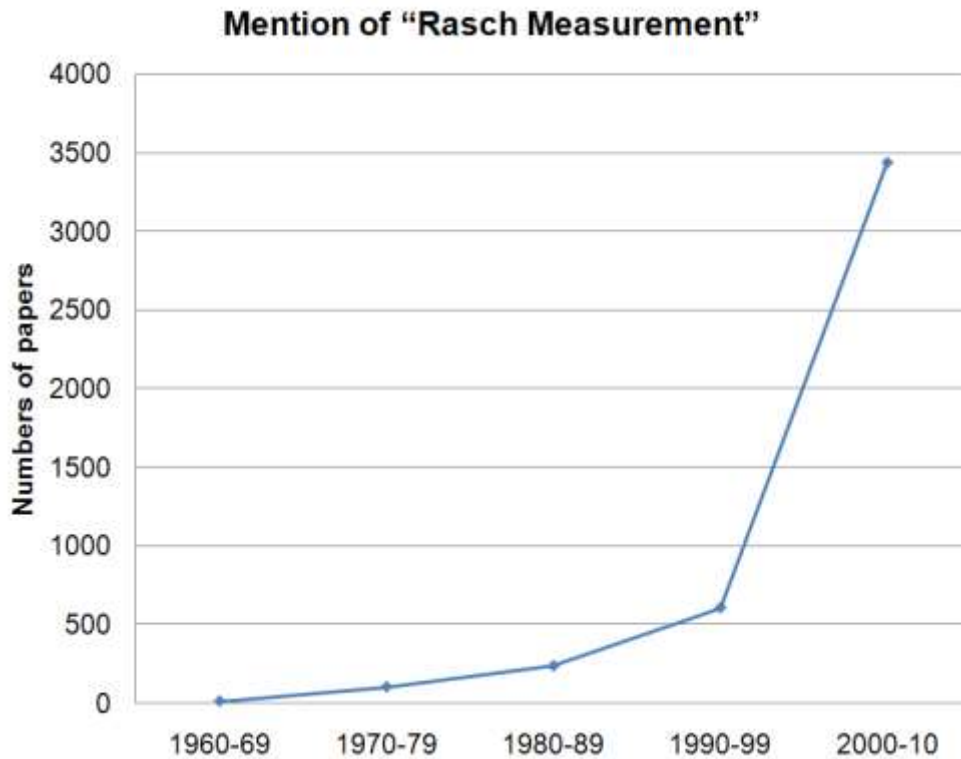
Goldstein writes: “the authors claim that there are no sample distributional assumptions associated with the Rasch model. This cannot be true, however, since the procedures used to estimate the model parameters, such as maximum likelihood, necessarily make distributional assumptions.”

But, as Andrich (2014) notes, “the conditional distribution of responses, (conditioned on the total score for each person - the sufficient statistic), is independent of the person parameters and therefore independent of the distribution of persons.” In other words the item estimates are person distribution free. Of course it is possible to make assumptions about distributions in making estimations but this is not necessary. (For further information see 2005, 2010)

Time warp

In the opening sentence of his response Goldstein that "Panayides et al. (2010) appear to be stuck in a time warp" (p. 4). But even though Rasch models were originally designed for use in educational assessment, over the last two decades there has been a remarkable increase in their use. Tymms (2013) illustrated this increase in an investigation into the number of times the phrase *Rasch measurement* has been used in published articles over the last five decades according to Google Scholar. The results of his enquiry are shown below.

Figure 1. Number of references to Rasch measurement in articles (1960-2010)



One wonders who is in the time warp.

Statistics or measurement?

Goldstein's sixth assertion is that we are requiring the data to fit the model rather than finding a model that fits the data. If the data do not fit the model, it indicates a potential flaw in the data. The problem with trying to find more and more complex models to fit imperfect data is that the interpretation of the results becomes more and more obscure. Goldstein looks for models that best describe the data at hand: models that can accommodate all peculiarities in the data regardless of whether they contribute to meaningful measurement. This should not be the basis of measurement. Researchers should ensure that the data conform to the principles of measurement before analysis.

"The Rasch model is a measurement abstraction which enables researchers in education and the social sciences to establish quantitative variables such as those commonly found in physical measurement. Where data fit the Rasch models the aforementioned properties are confirmed [continuous variables, need for linearity, equal repetitive units, scale magnitudes with the properties of real numbers,

precision and distribution-free measurement] and fundamental measurement follows" (Panayides, 2014, p. 9).

References

.

Andrich, D. (2005). Rasch models for ordered response categories. *Encyclopedia of Statistics in Behavioral Science*.

Andrich, D. (2010). Sufficiency and conditional estimation of person parameters in the polytomous Rasch model. *Psychometrika*, 75(2), 292-308.

Andrich, D. (2014) Personal communication

Goldstein H. (2010) Rasch measurement: a response to Payanides, Robinson and Tymms. www.bristol.ac.uk/cmm/team/hg/response-to-payanides.pdf

Hambleton, R. K., Swaminathan, H. Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. California: SAGE Publications, Inc.

Masters, G. N., and Keeves, J. P. (1999). *Advances in measurement in educational research and assessment*. Amsterdam: Pergamon.

Linacre, J. M. (1998a). Detecting multidimensionality: Which residual data-type works best? *Journal of Outcome Measurement*, 2(3), 266-283.

Linacre, J. M. & Fisher, W. P. (2012). Harvey Goldstein's Objections to Rasch Measurement: A response from Linacre and Fisher. *Rasch Measurement Transactions*, 2012, 26(3), 1383-1389. Available at <http://www.rasch.org/rmt/rmt263f.htm>

Payanides, P., Robinson, C. and Tymms, P. (2010). The assessment revolution that has passed England by: Rasch measurement. *British Educational Research Journal* 36: 611-626

Panayides, P. (2014). *Rasch Measurement in the Social Sciences and Quality of Life Research*. LAP Lambert Academic Publishing: Saarbrücken, Germany.

Smith Jr., E. V. (2004). Detecting and evaluating the impact of multidimensionality using item fit statistics and principal components analysis of residuals. In E. V Smith Jr, and R. M Smith (Eds), *Introduction to Rasch Measurement* (575-599). Minnesota: JAM Press.

Tymms, P. (2013). *Rasch: Some history and some possibilities* [PowerPoint slides]. Presentation at the 7th Rasch Day, Manchester (8th March 2013).

Wright, B. D., and Linacre, J. M. (1989). Observations are always ordinal; measurements, however, must be interval. *Archives of Physical Medicine and Rehabilitation*, 70, 857-860.
Available at: <http://www.rasch.org/memo44.htm>