REFERENCES

GOLDSTEIN, H. (1979) Consequences of using the Rasch model for educational assessment, *British Educational Research Journal*, 5, pp. 211–220.

PREECE, P.F.W. (1979) Objective measurement in education: the Rasch model, *School Science Review*, 60, pp. 770–773.

SHOEMAKER, D.M. (1975) Towards a framework for achievement testing, *Review of Educational Research*, 45, pp. 127–147.

A REJOINDER TO PREECE

H. GOLDSTEIN, *University of London*

Dr Preece's response to my article (Goldstein, 1979) on the Rasch model highlights a distinction between two different approaches to educational measurement.

He defends the 'unidimensionality' of a test as a means of conveying information about an underlying 'trait', and in so doing he is making two important assumptions. The first of these is that responses to test items are in fact determined solely (or at least overwhelmingly) by a single trait. Unless it is to be regarded as a matter of faith, such an assumption does require careful empirical verification. Moreover, there is considerable difficulty in deciding what is meant by a unidimensional trait. To define it in terms of test item 'content' requires strong theoretical assumptions linking item responses to individual knowledge and behaviour. This is difficult enough with even an apparently simple psychological test such as digit span (Jensen, 1969), let alone educational assessment procedures. The Rasch model adopts a different approach. It effectively defines a unidimensional set of items as those which happen to 'fit' the model. Unfortunately, items can 'fit' the model for all kinds of reasons, many of which may have little to do with unidimensionality of content. Thus model fitting can occur without reference to content, and one therefore cannot use the model of itself in order to establish content unidimensionality. The distinction between these two uses of the word unidimensionality seems to me to be very important and the weakness of Dr Preece's discussion lies precisely in his confusion of the two.

The second assumption is that educational measurement ought to be concerned with trait measurement. There are large areas, however, such as the public examinations system, where this is clearly not so and where some heterogeneous averaging of marks typically is required. Furthermore, I also argued in my article that there is no good reason to suppose even that tests of a narrowly defined attainment should be 'content' unidimensional. The fact that most test score theory does take unidimensionality for granted seems to me to reflect its *psychological* origins and the (possible reasonable) assumption that any quantitative theory for dealing with fundamental processes should attempt to develop unidimensional measuring instruments. My point is that *educational* measurement is generally quite a different activity, and my concern is that traditional psychometric theory may come to be used too automatically in an inappropriate context.

Dr Preece's clock analogy does, I think, somewhat miss the point I was making. As he says, it is perfectly possible for physical measuring instruments such as clocks to go out of calibration and a physicist would wish to find explanations for this. Thus the comparison, say, of atomic and astronomical clocks might be an

appropriate area of investigation. Nevertheless, the fact that clocks may not keep in step does not in any way invalidate the notion of a clock, even though it may well lead us to abandon the idea of defining a time-scale simply as an average over a set of clocks. In the case of test items it may be interesting to study the reasons for difficulties which apparently change over time (or across curricula), but that would serve only to underline the logical contradiction inherent in trying to define time (or curriculum) invariant Rasch scales in education.

Correspondence: Professor H. Goldstein, University of London Institute of Education, Bedford Way, London WC1H 0AL.

REFERENCES

GOLDSTEIN, H. (1979) Consequences of using the Rasch model for educational assessment, *British Educational Research Journal*, 5, pp. 211–220.

JENSEN, A.R. (1969) How much can we boost IQ and scholastic achievement? *Harvard Educational Review*, 39, pp. 1–123.

PREECE. P.F.W. (1980) On rashly rejecting Rasch: a response to Goldstein, *British Educational Research Journal*, 6, pp. 209–211.