Preface

This is the second volume in the series, *Assessment and Learning,* from the Assessment & Support Team (formerly the Basic Competency Assessment Team). Compared to the previous publication this contains more papers from internationally known authors and it says much for their growing reputation in the field of assessment that team is able to call on such distinguished contributors. Teachers may not find some papers an easy read but if they persevere they will be rewarded by many rich insights into the increasingly complex world of educational assessment.

Nearly half the papers deal with student weaknesses in Languages and Mathematics and the means of identifying and dealing with these. One particularly provocative piece concerning extended writing asks whether it is always necessary to pick out and correct all errors or only the consistent ones. On more general assessment issues, two other strong themes emerge from the collection. In the first instance, the argument is put forward that for Assessment *of* Learning empirically constructed tests (the kind that require initial piloting , checking for internal consistency and item facility etc.) are time-consuming, costly to develop and can be problematic in terms of their predictive validity. It is suggested that in their place 'theory driven' tests, such as those based on the Rasch model, should become more readily available to teachers for use as both a formative and summative tool. Second, the need to extend the notion *of* Assessment *for* Learning (AfL) to embrace Assessment *as* Learning (AaL) is emphasised by several contributors based on the research evidence that feedback which promotes *forward* thinking by students pays rich dividends in future attainment gains.

There is much to be said for utilising a linear measure of ability which is dependent only on the difficulty of the item, in much the same way that in an old fashioned mercury thermometer the temperature of an object is proportional to the height of the liquid column. But there is no corresponding simple visual

effect in the use of Rasch scaling. One paper, for example, includes tables where the ability scales have minus scores and this concept may not be easy to understand when presented, for example, to the typical primary teacher. In an age of computer graphics that can create extraordinary 3-D representations it should be possible to produce pictorial images which simplify the presentation and the analysis for those classroom practitioners, who may lack the necessary mathematical ability to understand and therefore interpret the results in their present form. On the other hand, it could be argued that having teachers base their judgements about a student's ability on procedures they do not fully understand can set a dangerous precedent. A parallel exists in the field of educational research where it is possible for anyone who can construct a data matrix using Excel to carry out sophisticated statistical analyses using a package such as SPSS without any understanding of the underlying assumptions of the procedures and their limitations and sometimes with disastrous consequences. It is my hope that future issues will pursue this debate further.

Another issue where further research on theory driven tests is needed concerns construct validity, particularly if the analysis of data using the Rasch methodology is to be used for diagnostic purposes. At the moment there is a tendency for those who construct such tests to rely heavily on face validity and it is sometimes not easy to discern why some items fit on a uni-dimensional scale while others do not. Many years ago I was involved in a project attempting to measure the effects of different teaching styles in science on student attainment[1]. Our hypothesis was that teachers who adopted an enquiry approach, as opposed to a didactic problem solving one, would enhance students' scientific reasoning. We therefore set about constructing tests designed to measure higher order skills such interpreting data, making inferences from data, and formulating hypotheses. But to employ these skills required students to have an understanding of the theories behind the problems we set in the written tests, so we were unsure when a student gave a wrong answer whether it

[1] Eggleston, J., Galton, M. and Jones, M. (1976). *Processes and products of science teaching*. London: Macmillan (for the Schools Council).

was because s/he lacked a theoretical grasp of the principle or whether they were unable to apply the theory to the particular question. One of us suggested we should therefore include a brief account of the theory in the stem of the question which we did, for example, in one case providing a brief summary of the laws of reflection. But then we discovered that our consistency measure was distinguishing between the candidates' ability to comprehend the content of the question stem and not their problem solving skill.

In such cases the answer lies in interrogating the students further in order to discover how they arrived at their respective answers and in time helping them to conduct their own self-evaluations. This form of questioning lies at heart of both AfL and even more importantly AaL and raises the question of where our efforts to reform current assessment practice should be directed in the immediate future. Is it to be educating teachers to make regular use of procedures such as Rasch for formative testing purposes or should we concentrate on the immense task of changing the public (and official) perceptions as to the key purposes of assessment, particularly its use in AfL and AaL. Only by doing this will we remove the pressure on teachers to continue to use conventional testing methods. One contributor argues, and I agree, that this will require all involved in education at the grass root level to wage a 'quiet' but 'determined' revolution.

This collection of papers and the previous volume in the series are a vital component of this revolutionary process. I look forward to more of the same in future issues, for in preparing the next generations of students to face the demands of the twenty-first century there is no more important task than to convince a somewhat sceptical public that assessment involves more than setting regular written tests.

Maurice Galton