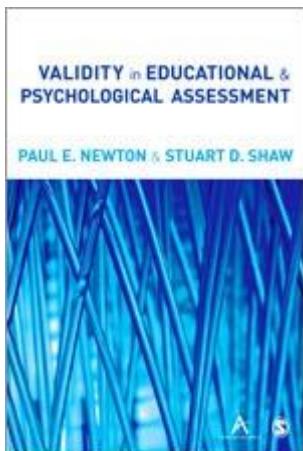


Book Review



Validity in Educational and Psychological Assessment

By **Paul Newton** and **Stuart Shaw**

London: Sage Publications (2014)

Hardcover £87.00 (ISBN: 9781446253229)

Paperback: £27.99 (ISBN 9781446253236)

Reviewed by **Lydia May Townsend**

This book provides a step-by-step guide to understanding the complexities of validity in assessment. Split into easily digestible chunks, it navigates the history and developments in validity, starting in the mid-1800's and ending up in the modern day. Largely situated in western understandings of validity, it links these histories to the wider educational contexts of the time and it makes a particular point of unpicking the historic ideas of validity in assessment that still dominate modern understandings. The particular strength of this book is to anticipate the questions that will naturally arise in the reader's mind and answer them in a carefully structured manner. It also need not be read as a whole, though it is a very interesting read if you do have the time. Rather, the reader can dip in and out, focussing on those areas which are of most interest without losing the general thread of the arguments.

The book begins by detailing what exactly is meant by validity. It demonstrates how one might define validity, and how this understanding differs when validity is used in research or in measurement. The crux of the argument centres on the assertion that 'validity for measurement is relevant only for conclusions that relate specifically to measurement' (pg 26). For those familiar with research in this field this conclusion may come as no surprise, but for those with no prior knowledge the authors give a brief explanation of why this must be so. The first chapter proves a comfort to the confused reader, who may have already unsuccessfully attempted to get to grips with validity. It demonstrates the sheer number of different methods that have been employed to try to establish validity in assessment. It then goes on to describe the different terms that are used in conversations of validity, succeeding in being succinct and clear in its approach.

What may come as a surprise is how relatively recent are conceptions of validity. It may be reasonable to assume that validity has always been a human concern, however the book dispels this myth. In Chapter 2 it takes the reader on a journey through the 1800's to the start of the 1950's, focussing predominantly on the early to mid-part of the 20th century. It is striking to see quite how influential western university systems were in shaping modern ideas of assessment still familiar today. If there can be one criticism of this book, it is in the brevity of its exploration of assessment before the 1920's. It makes no mention, for example, of the well-established assessment systems of the Chinese civil service which started in Sui dynasty (581-618 CE). This criticism is minor, since the authors prewarn the reader that this section will be short, and it does not detract from the overall conclusions of the book.

The third chapter brings the reader into the period 1952-1974, and details parts of assessment with which many readers will find familiar from their own education. Moving on from the array of approaches which characterised the earlier part of the 20th century, and the problems such

an array caused, we are drawn into an era of standards. Here we see the terms content validity, predictive validity, concurrent validity and construct validity described in relation to their meaning at the time. We are then introduced to Cronbach and move seamlessly into the fourth chapter, which details ideas of validity from 1975-1999. Here we are introduced to Messick and ideas of a matrix in validity. By both examining why the development of the matrix moved validity understanding forward, and subsequently criticising its approach, the reader is given a detailed insight of how academic ideas fall in and out of favour. This is particularly useful to the student who is still under the impression that validity is a singular and absolute concept, for this chapter dispels that certainty.

The final two chapters (chapters five and six) take the reader through validity in the new millennia and bring us to modern thoughts. We see the desire to simplify validity, battling with difficulties such simplification holds. The authors make a point as the book draws to an end in positioning the modern ideas within the historic context which they used the preceding chapters to establish. Here the explanations become meatier and less easy for the novice to follow. The ideas of the matrix explored are entangled with ideas of policy. If the reader can get through this necessary complexity, they will find this last chapter particularly insightful and useful for understanding how ideas of validity may develop as we move further into the 21st century.

This book would be particularly useful to any persons charged with creating or implementing tests. This includes higher education students, teachers, policy makers, psychologists, all of whom have some vested interested in assessment. Though academic in tone, it makes a point of explaining ideas in the simplest possible way and does not require the reader to have much prior knowledge of assessment at all. It is also of great use to those among the academic community who wish to understand why they must contend with the assessment system we have, and who would like a strong foundation on which to challenge the *status quo*.