



UvA-DARE (Digital Academic Repository)

Conceptual issues in psychological measurement

Borsboom, D.

Publication date
2003

[Link to publication](#)

Citation for published version (APA):

Borsboom, D. (2003). *Conceptual issues in psychological measurement*. [Thesis, fully internal, Universiteit van Amsterdam].

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

PREFACE

In hindsight, it is notoriously difficult to escape the impression that there have been certain turning points, or landmark events, in one's intellectual development. In my personal reconstruction of this process, two such events occurred when I was still a student, and both involved my later supervisors.

The first took place when I had just enrolled as a student at the Psychological Methods department. Like all new students in methodology, I followed a course on formal models which was given by Don Mellenbergh. Don introduced the classical test theory model by stating that he was going to do something strange. He was going to brainwash people inbetween testing occasions. When he then took the expectation over such replications, and subsequently introduced a population sampling scheme, he could drop the reference to the expectation of replications again, and all kinds of elegant expressions followed for important concepts like reliability. This seemed almost like a magic trick to me. It was the first time I started suspecting that there was something strange with the formal models that we use in the analysis of psychological test scores, and that it had to do with the application of the expectation operator.

The second event concerned a question posed by Jaap van Heerden at a technical colloquium about item bias. The speaker casually remarked that the latent variable could be considered the cause of its indicators. After the talk, Jaap looked over his glasses in a fashion that I later learned to be very characteristic of him, and said, "so, you would say that mentalistic concepts can play a causal role. . .". Then there was silence. When the speaker had regained his calm, he tried to defend his position by invoking the analogous thesis that mass was obviously the cause of readings on a balance scale. Jaap turned out to doubt this too. I remember thinking, "that man is crazy", but the question did not let go of me. I do not know whether Jaap's degree of inquisitiveness is to be considered pathological, but if it is, then I am afraid that I have developed a similar syndrom over time, so I hope not.

The combination of these two fundamental problems – that is, the interpretation of probability in psychometrics, and the theoretical status of concepts like latent variables – are the basis of this book. For when I became a Ph D. student with Don and Jaap, it seemed like a good idea to devote some attention to these issues. However, in searching the literature, I failed to find a thorough analysis of these problems, although possible conceptualizations were sometimes alluded to – usually in footnotes, or in paragraphs that were accompanied by notes like "reading of this section can be omitted without loss of continuity". This surprised me. I did find a wealth of literature on related questions in the philosophy of science,

and I also found many psychological articles on the status of substantive traits, like general intelligence and the Big Five. None of these papers, however, was directly concerned with latent variable theory as it is used in psychometrics. Applying the frameworks of philosophy of science to the theoretical status of latent variables was, in my opinion, very clarifying, although the resulting paper (Chapter 3 in the present dissertation) can arguably be said to raise more problems than it solves. It turned out, however, that to apply philosophy of science to the question how technical concepts in psychometrics may relate to substantive concepts in psychology, was a very good way to elucidate the issues involved. This dissertation applies this line of thinking to different theories of psychological measurement. I have chosen to include the true score model, the latent variable model, and the representational measurement model in the analysis. The treatment must therefore be considered incomplete. Missing from the analysis are generalizability theory and multidimensional scaling. Also missing are chapters about the theoretical status of observed scores, and about a mysterious species of entities known as 'constructs'. Like life itself, however, a Ph. D. studentship is short, and choices have to be made.

Although the material included in this dissertation should be read as a book, the chapters are based on papers that have been published elsewhere or are currently submitted. Specifically, Chapter 2 is partly based on Borsboom & Mellenbergh (2002). Chapter 3 is a slightly adapted version of Borsboom, Mellenbergh, & Van Heerden (*in press*). Parts of Chapter 4 are based on Borsboom & Mellenbergh (*submitted*). The last part of Chapter 5 is based on Borsboom, Van Heerden, & Mellenbergh (*in press*); I must, however, say that I no longer subscribe to the conclusion of that paper, and my views on validity are now better represented in Chapter 6, which is based on Borsboom, Mellenbergh, & Van Heerden (*submitted*). An overview of and introduction to this material is given in the first chapter of the book. In addition, I have chosen to include two articles (Borsboom, Mellenbergh, & Van Heerden, 2002-a; Borsboom, Mellenbergh, & Van Heerden, 2002-b) as appendices.