

# Preface

In this edited volume, 15 research syntheses of the validity evidence reported in different research areas are presented. The chapters were purposefully chosen to reflect a wide variety of disciplines, journals, or measures. Eight of the chapters focused on particular journals ranging from measurement and assessment journals like *Educational and Psychological Measurement*, *Psychological Assessment*, to international counterparts such as the *European Journal of Psychological Assessment*, as well as *Social Indicators Research: An International and Interdisciplinary Journal for Quality of Life Measurement*. In total 11 journals in a variety of disciplines that were North American, European, or International focused were surveyed in the chapters. From these journals one can see the far reach that we aimed to contain. Likewise, nine chapters focused on key tests, measures, or assessment tools that provide a sense of validation practices in particular areas of assessment. Note that one chapter focused both on a group of journals as well as particular measures. In short, in essence, we are studying the scholarly genre of validation reports and how this genre frames validity theory and practices.

Each chapter is meant to stand alone and hence one could read a sub-set of the chapters in any order. The “free-standing” nature of the chapters is important because readers may want to focus on one, or more chapters, because of the vast array of domains, topics, and measures we covered.

We were mindful that we wanted each chapter to be both unique but also use some common framework. Therefore, we decided that all chapters would, at least, follow the generic framework in the *Standards* (AERA et al. 1999) wherein five sources of validity evidence were of focus: (a) content-related, (b) response processes, (c) internal structure, (d) associations with other variables, and (e) consequences. The syntheses also addressed whether recent work in validity theory was cited as informing the validation practice (e.g., Hubley and Zumbo 1996, 2011, 2013; Kane 2006; Messick 1989; Zumbo 2007, 2009).

This volume represents a broad sampling of educational, psychosocial, and health research settings, giving us an extensive evidential basis to build upon earlier studies by Cizek and his colleagues (2008, 2010). It is worth noting that the chapters in this volume commonly used a sampling of papers because unlike Cizek et al. (2010) who

used a word search and hence were able to include hundreds of papers, the authors herein based their synthesis on a close read of the papers and not an automated word search. Therefore, in our authors' cases, the number of papers is limited by the methodology. This methodology has the benefit of contextualizing the findings reported in each of the papers being synthesized, and overall there are hundreds of papers (more than 500) reviewed in detail.

We would like to outline for you the general principles and ethos of the book. The book is organized in five parts. Part I consists of an introductory chapter that sets the stage for and purposes of the book, and a second chapter reviewing standards and guidelines for validation practices in a variety of academic disciplines and jurisdictions. Part II includes three chapters devoted to quality of life, wellbeing, and life satisfaction. Part III consists of six chapters broadly reflecting psychology and education. Part IV consists of six chapters in the broad domains of health and medicine, including health psychology, patient-reported outcomes, and medical education. It should be noted that the chapters in Parts II–IV overlap a great deal in focus (which is not surprising given the overall purpose of the book) and could be re-arranged with different section headings. The closing part includes two concluding chapters. The first is a “meta-synthesis” of the 15 research syntheses and the closing chapter takes the reader back to the broad focus of the whole volume.

Because of its breadth of scope and purpose, this book is a high watermark in the history of measurement, testing, and assessment because it documents what people do when they validate their tests, measures, or assessment instruments in a wide variety of disciplines and regions of the world. This focus on validation practices is interesting in and of itself and will influence both future validation studies and theorizing in validity. In part, it documents how validity theory is influencing validation practices, and it also guides us in developing a plan for validation work. In broad terms, we aimed to answer the question: What passes as validity evidence? In other words, when people validate a measure, what do they do? What does the academic community accept as evidence of measurement validity in its scholarly journals? It is important to note that our focus was not on whether the score inferences drawn from any particular measure, test, or assessment are “valid” but rather on the sources and kinds of validity evidence that are reported in the published research literature.

Like all studies, there are limitations to our work; the largest one is by design. Our focus is on papers published in scholarly journals. We did not include any synthesis of what testing organizations, testing companies, or professional test publishers are doing in their validation practices as reflected in test manuals or validation studies within their organizations. Some of this is captured in the work of Cizek and his colleagues (2008) in their study of the *Mental Measurement Year-book*<sup>1</sup>; however, some of this information is also difficult to obtain because several testing organizations treat their validation studies as propriety information. As a reminder, however, our focus was on papers published in scholarly journals, and as

---

<sup>1</sup> Curiously, their overall findings are consistent with ours.

we show in our search of the PsycInfo database in Figs. 1.1, 1.2, and 1.3, we have a large body of work to select from and hence our focus is warranted.

We would like to close by acknowledging the impressive body of work that our collaborators amassed. To support the reading of each chapter, each chapter author was asked to speak to validity theory in their domain and, where possible, make recommendations for validation practices. There is much gold to be mined for validity theorists and practitioners in the closing sections of each chapter. In addition to our own review of each of the chapters, we would like to thank Dr. Katie Gunnell, Dr. Rebecca (Beck) Collie, Michelle (Yue) Chen, and Dr. Dallie Sandilands who each provided valuable feedback for several chapters.

Vancouver, BC, Canada

Bruno D. Zumbo

Eric K.H. Chan

## References

- American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Cizek, G. J., Rosenberg, S. L., & Koons, H. H. (2008). Sources of validity evidence for educational and psychological tests. *Educational and Psychological Measurement*, 68, 397–412.
- Cizek, G. J., Bowen, D., & Church, K. (2010). Sources of validity evidence for educational and psychological tests: A follow-up study. *Educational and Psychological Measurement*, 70, 732–743.
- Hubley, A. M., & Zumbo, B. D. (1996). A dialectic on validity: Where we have been and where we are going. *The Journal of General Psychology*, 123, 207–215.
- Hubley, A. M., & Zumbo, B. D. (2011). Validity and the consequences of test interpretation and use. *Social Indicators Research*, 103, 219–230.
- Hubley, A. M., & Zumbo, B. D. (2013). Psychometric characteristics of assessment procedures: An overview. In K. F. Geisinger (Ed.), *APA handbook of testing and assessment in psychology* (Vol. 1, pp. 3–19). Washington, DC: American Psychological Association Press.
- Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 17–64). Westport: American Council on Education/Praeger.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). New York: American Council on Education and Macmillan.
- Zumbo, B. D. (2007). Validity: Foundational issues and statistical methodology. In C. R. Rao & S. Sinharay (Eds.), *Psychometrics*, Handbook of statistics (Vol. 26, pp. 45–79). Amsterdam: Elsevier.
- Zumbo, B. D. (2009). Validity as contextualized and pragmatic explanation, and its implications for validation practice. In R. W. Lissitz (Ed.), *The concept of validity: Revisions, new directions and applications* (pp. 65–82). Charlotte: IAP – Information Age Publishing.

Validity and Validation in Social, Behavioral, and Health  
Sciences

Zumbo, B.D.; Chan, E.K.H. (Eds.)

2014, XV, 327 p. 15 illus., 2 illus. in color., Hardcover

ISBN: 978-3-319-07793-2