Book Review

Jennifer Edgar¹

Paul C. Beatty, Debbie Collins, Lyn Kaye, Jose-Luis Padilla, Gordon B. Willis, and Amanda Wilmot. *Advances in Questionnaire Design, Development, Evaluation and Testing*. 2019, Wiley, ISBN: 978-1-119-26362-3, 816 pages.

Researchers have been thinking about, and researching the most effective approaches to do questionnaire design, development, evaluation and testing for a long time. Since well before the original International Conference on Questionnaire Design, Development, Evaluation and Testing (QDET) was held in 2002, the quest to find the most effective and efficient approaches to collect survey data has been the focus of researchers in academia, government agencies and private companies across the globe. In 2016, many of these researchers came together to share findings and innovations, and readers are fortunate that the information shared at that second conference (QDET2) was pulled together into this book to document the advances in the field, as well as the thoughts on questions yet to be answered by some of the most prominent players in the field.

Advances in Questionnaire Design, Development, Evaluation and Testing starts as one would hope, laying out the current state of affairs and highlighting issues most likely to face the field in the future. The first chapter was a perfect introduction for the book. Willis provided historical context while looking into the future to foreshadow several of the upcoming chapters. Dillman stayed at the big-picture level in the second chapter, connecting QDET2 back to the original QDET by identifying some of the key issues speakers tackled. Readers may be struck by the number of issues that remain relevant today, and are addressed explicitly in the volume.

Three additional introductory chapters lay out the current thinking around questionnaire design and evaluation, painting a rich picture of what we know after all the years of questionnaire design and evaluation research and pointing out the areas we all still need to work through. It's certainly difficult to argue with authors such as Willis, Tourangeau or Dillman based on their status in the field; they and their co-authors do an excellent job laying out the current state of affairs.

Throughout Part One of the book, there is a careful balancing of optimism with how far the field has progressed (e.g., Dillman no longer feeling as frustrated with the state of questionnaire evaluation as he was after the first QDET) with caution (e.g., Willis noting that there is still work to be done before we can speak to "what is a good question"). These overview chapters provide the audience with a useful perspective with which to view the subsequent chapters which tackle specific design, evaluation and testing topics.

¹ Bureau of Labor Statistics, Office of Survey Methods Research, 2 Massachusetts Ave., NE Washington D.C., U.S.A. Email: edgar.jennifer@bls.gov

In response to the initial chapters of the book, I'd like to challenge both Part I authors, and readers, to think about that question. If, after more than 20 years, evaluating the quality of a question, or a survey is elusive, what does that mean? Perhaps, the goal is not to identify a 'good' question per se, but instead to identify potential problems with the question, and then re-evaluate to determine if we reduced the likelihood of that problem occurring. Maybe we'll never get to a set of known 'good' questions, but instead can have a set of evaluation and testing techniques to get us to 'better.' Given the number of high quality studies looking at the impact of questionnaire design documented in this book, it is clear that researchers find design and evaluation efforts worthwhile, even if we're not necessarily able to quantify the improvement in question quality. While many of us would appreciate the opportunity to conduct research to fully evaluate our evaluation methods, in the current climate of declining response rates and other challenges, researchers are limited to addressing immediate problems like how to present a grid question on a mobile phone (Dale and Walsoe) or how to collect consistent information across cultures and languages (Smith). And, in the case of the authors represented in this volume, it is clear that they are tackling their real-world issues in a way that aims to address the overarching goal of administering high quality questionnaires.

The remaining chapters in this book provide readers information on a wide variety of subjects, and will undoubtedly serve as a valuable reference. Some chapters provide a useful overview of a topic, such as Yan and colleagues who provide a thorough overview of respondent burden including the relevant literature before adding some empirical evidence about which survey characteristics are predictors of burden. Others are more targeted, Nichols et al. who give readers a behind the curtain view of usability testing at the U.S. Census Bureau.

There are also articles that take a step back and provide insight into conducting and managing questionnaire evaluation research such as Stapelton et al. who talk about complex cognitive testing projects and Jans et al. who share an overview of how they leveraged multiple methods and iterative testing and even offer suggestions for future studies considering doing the same.

As a whole, the chapters cover a good mix of theory and research and provide readers a valued resource on a variety of topics. For readers who were not fortunate to attend QDET2, this volume provides a sense of the breadth and depth of questionnaire evaluation topics covered.

Reflecting on this book may leave the reader pondering "What issues will QDET3 tackle?" Will the Q no longer stand for questionnaire, instead representing the focus on evaluating Quality of blended data, administrative data, non-designed data, etc.? Or will we still be struggling with how to present questions to respondents to collect the most accurate information we can with minimal burden? While personally I am confident that surveys are here to stay, there are certainly opportunities to look at the techniques such as those presented in this book, that have been fine-tuned over the decades and identify ways to apply them to new types of information collections. As long as we seek to draw conclusions about people, regardless if it's by asking respondents directly or indirectly through non-designed data, we will always need to consider and evaluate the quality of information collected.

In summary, this book should be on the shelf of anyone who is actively working in the field. Not only does it provide survey practitioners with empirical information on a wide range of topics (e.g., measuring disability equality), it also provides thought-provoking information about methodology in the field (e.g., online pretesting methods or cross-cultural surveys). The combination of the two will serve readers well, regardless of their experience in the field, both novices and experts will find much to learn here. The topics and issues presented here, both those that seem resolved and the questions raised but not yet answered will undoubtedly stay relevant as we all continue to seek the most effective and efficient ways to design, develop, evaluate and test our questionnaires in the years to come.