

INTRODUCTION: WHY IS THIS BOOK NECESSARY?

MATTHEW C. MAKEL AND JONATHAN A. PLUCKER

The preamble to the U.S. Constitution contains the phrase “in Order to form a more perfect Union.” This sentiment acknowledges that perfection could likely not be obtained (and perhaps not even wholly understood). We approach this book from much the same mind-set; we do not know what a perfect psychological science looks like, but we believe that scholars should and can strive to reach this goal. Our path to this book started in 2007, when one of us (Jonathan A. Plucker) dropped a page from a magazine on the other’s desk and said, “We need to think about stuff like this.” The page was a summary of some new findings showing that, in marketing research, studies were rarely replicated and, when they were, the replicating studies often failed to confirm the original findings (Evanschitzky, Baumgarth, Hubbard, & Armstrong, 2007). And with that quick drop-in, a whole new world was opened to us.

<http://dx.doi.org/10.1037/0000033-001>

Toward a More Perfect Psychology: Improving Trust, Accuracy, and Transparency in Research, M. C. Makel and J. A. Plucker (Editors)

Copyright © 2017 by the American Psychological Association. All rights reserved.

The idea of needing to assess and improve the quality (not to mention the public image) of research was certainly not new to us. At the time, the federal Institute of Education Sciences had recently been created in response to perceptions of low research quality in education research. We initially focused our attention on gifted education and talent development research because it often relies on either general education findings or research conducted with small (and often idiosyncratic) samples of gifted students. As such, professionals in the field are often unsure whether and when research findings found in one context for one group of students can be generalized to other contexts and other groups of students. Our concerns fell quite nicely into two categories: issues pertaining to (a) direct versus (b) conceptual replications (see Schmidt, 2009, and Chapter 14, this volume).

The more we dug, the more we found that many others were also worried about issues of replicability and generalizability. One article that garnered perhaps the most attention had the provocative title “Why Most Published Research Findings Are False” (Ioannidis, 2005). This spurred us on, and we eventually learned that in attempts to replicate findings in medical research, researchers succeeded in replicating only six of 53 (11%) highly cited cancer trial studies (Begley & Ellis, 2012).

With such damaging findings coming from fields typically associated with higher levels of methodological rigor than the social sciences, our focus soon moved to psychology, and we began our first large-scale examination of replication in the social sciences, specifically within psychological research (Makel, Plucker, & Hegarty, 2012). Bem (2011)¹ published a controversial study on the existence of *psi* (e.g., telepathy) that served as a focal point for much ensuing conversation. However, the topic was only the tip of the controversy; its methods also spurred strong and immediate negative reaction (see, e.g., Sutton, 2012; Wagenmakers, Wetzels, Borsboom, & van der Maas, 2011). Bem’s provocative article certainly did not invent concerns about the psychology research process; indeed, they had been around for decades (e.g., Bakan, 1966; Cohen, 1994; Lykken, 1968; Rosenthal, 1966, 1979). However, we believe that the publicity the article received helped blaze the path for many of the new movements in psychology and the social sciences.

¹In the interest of full disclosure, we should share that one of us (Matthew C. Makel) took a class from Bem while earning a master’s degree. Bem was a wonderful teacher. In his class and his writing, Bem told the story of how he was brought into *psi* research as an outsider who could help the field with its methodological rigor. He was impressed with their methods and helped publish findings in a top psychology journal (Bem & Honorton, 1994, but cf. Hyman, 1994) because he knew how to navigate the publication process. In perhaps an unintentional irony, taking advantage of this expertise in how the publication system worked helped catalyze the current revolution in assessing research quality.

Nothing is new under the sun (e.g., Makel & Plucker, 2014). In 1984, Furchtgott made the following lament:

It would seem that after more than 30 years the [American Psychological Association] Publication Committee or an ad hoc committee should examine the publication policy pertaining to replications. Not only will this have an impact on investigations that are undertaken, but it will reduce the space devoted to the repetitious pleas to replicate experiments. (p. 1316)

In that spirit, with this book we hope to take a few steps backward (and perhaps one or two to the side) so that the field may accelerate its forward momentum, not just on replication but also on other relevant issues. At its foundational level, the heart of science is that its methods allow for others to believe its results. This foundation is served by trust, accuracy, and transparency. It is only with all three that science succeeds. Without one, the others are weakened.

We also seek to introduce readers to the many proposals that are being developed to help strengthen the field's research. It is our hope that this book will help researchers not just as they develop their own research ideas, question, methods, and analyses but also as they evaluate and respond to the research of others.

ORGANIZATION OF THE BOOK

To develop and discuss these concepts (and a great deal more) in more detail, we have organized this book into five sections to help highlight the different perspectives researchers take throughout the research process as well as their careers.

Part I has the central theme of asking what individual researchers can do to increase research quality while focusing on different steps of the research process. Chapter 1, by Oshin Vartanian, discusses three factors that increase the probability that scientists will pursue important phenomena and processes: the researcher's (a) choice of theory, (b) pursuit of cumulative science, and (c) explicit focus on problem finding. These processes often occur (although not often enough) before data collection, at the idea generation stage. Chapter 2, by Alison Ledgerwood, Courtney K. Soderberg, and Jehan Sparks, covers how to design a study to maximize its informational value. In Chapter 3, Matthew T. McBee and Samuel H. Field focus on methodological issues pertaining to confirmatory study design and data analysis techniques. The goal is to make sure that researchers remain focused on results that matter.

The remaining chapters in Part I broaden the scope to discuss how practices conducted for individual studies influence the larger research community. Chapter 4, by Terri D. Pigott, Ryan T. Williams, and Jeffrey C. Valentine, examines selective outcome reporting. Research studies are not portfolios where researchers can manipulate post hoc what they asked, what they analyzed, and what they found. As the chapter authors discuss, when such manipulation happens researchers paint an inaccurate portrait of what has been evaluated, what works, and what should be done. Alyson L. Lavigne and Thomas L. Good conclude Part I with Chapter 5, in which they explain how citations serve as a researcher's intellectual footprint. They focus in particular on who is cited, why, and what it means to be cited (or not).

Whereas Part I focuses on different aspects of the research process, Part II provides perspectives on different roles researchers take as part of the research process (i.e., peer reviewers, disseminators of research to the public, and disseminators of research to other researchers). In their careers, researchers often have responsibilities beyond merely conducting research, including connecting research to practice, reviewing the research of others, and disseminating research. Part II begins with Chapter 6, by Jennifer L. Richler and Isabel Gauthier. Drawing on their considerable editorial experience, these authors provide advice to authors, peer reviewers, and editors about all stages of the paper submission process. In Chapter 7, Howard Gardner discusses how one should communicate research to the public in ways outside of typical academic publishing. Chapter 8, by Jonathan A. Plucker and Paul J. Silvia, addresses how to disseminate one's research in ways that maximize its influence on the field. Together, the chapters in Part II help readers understand what can be done with academic research and contextualize how others will apply, consume, and share research results.

The chapters in Part III provide some views from the field regarding research quality. The authors in this section offer examples of how different subfields can have unique strengths and struggles with producing rigorous research results. Chapter 9, by Mark Berends and Megan J. Austin, discusses forming positive relationships with practitioners to improve research quality while enhancing the likelihood that research findings will be implemented by practitioners. Amy Lynne Shelton, author of Chapter 10, addresses research in cognitive neuroscience. She highlights the importance of maintaining methodological rigor while harnessing technological sophistication to answer relevant questions. In Chapter 11, Brian M. D'Onofrio, Richard J. Viken, and William P. Hetrick focus on the importance of grounding clinical psychology in scientific epistemology, taking advantage of the methods and knowledge in related fields and integrating research with clinical practice in order to make both as reproducible as possible. Finally, Chapter 12, which we coauthored,

introduces readers to some potential ethical pitfalls they may experience in careers as professional researchers.

Part IV of the book focuses on three central components of reproducibility of research: (a) open data, (b) replication, and (c) meta-analysis. In Chapter 13, Jelte M. Wicherts addresses data reanalysis and open data, two emerging practices that help provide a form of “sunshine law” that allows others to verify specific results, thus increasing the accuracy and trust of social science results. Another method of accomplishing this that has garnered growing attention is replication, covered in Chapter 14 by Stefan Schmidt. Purposefully repeating the same methods of a previous researchers or systematically modifying small components of previous research have long been desired but only recently have begun to be perceived more positively within psychology research community. Chapter 15, by Ryan T. Williams, Joshua R. Polanin, and Terri D. Pigott, focuses on meta-analysis and reproducibility. Systematic reviews and meta-analyses have long served as a ballast of the field by removing the idiosyncratic blips that are often associated with individual studies. Together, the chapters of Part IV cover field-wide practices whose prevalence can help make psychology a stronger, more rigorous science.

A book about research quality and reproducibility certainly should not rely on a single perspective of “what it all means.” Therefore, the final section, Part V, consists of two synthesis chapters: one by Jeffrey K. Smith, Lisa F. Smith, and Benjamin K. Smith (Chapter 16), and one coauthored by us (Chapter 17). Only with time will we know which trends will evolve into common practice and which will turn out to be fads. However, instead of simply throwing our hands into the air and waiting, each team of authors assesses which of the discussed practices will most help the field produce more accurate, trustworthy, and reproducible results.

In our invitations to the chapter authors we explained that our goal for this book was to help researchers improve the quality of their own work as well as that of colleagues. Our perspective is largely based on our own studies of replication in the social sciences, but we expanded the focus to address research quality in the social sciences in general. We told authors that our broad vision of each chapter would be to introduce the reader to many of these events and issues and to discuss how the field is (and should be) responding. Our priority with this book is to share the evolving ideas of how to create a field that yields trustworthy, reproducible, and accurate research findings to a broad audience of social scientists as well to future researchers so that their training includes new practices of quality research methods, not just the status quo approaches.

The ideas presented in this volume will not perfect science. Solutions to one set of problems can often create new, unintended consequences. With the

creation of processes like meta-analysis, multilevel modeling, and regression discontinuity, our statistical and methodological sophistication has grown tremendously in the past 50 years. Also, with the flood of new movements seeking to improve how science is conducted (e.g., open data, preregistration), we expect that researchers 50 years from now will be conducting even better science we are today.

From this perspective, we are cautiously optimistic. As is often recommended in recovery programs, psychology has begun to admit it has a problem, and a great many psychologists have been working diligently to develop ways to minimize, remove, work around, and avoid many of the problems that plague the field. We hope books like this one help accelerate the field's path toward increasing scientific quality and rigor and thus a more perfect psychology.

REFERENCES

- Bakan, D. (1966). The test of significance in psychological research. *Psychological Bulletin*, 66, 423–437. <http://dx.doi.org/10.1037/h0020412>
- Begley, C. G., & Ellis, L. M. (2012, March 29). Drug development: Raise standards for preclinical cancer research. *Nature*, 483, 531–533. <http://dx.doi.org/10.1038/483531a>
- Bem, D. J. (2011). Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology*, 100, 407–425. <http://dx.doi.org/10.1037/a0021524>
- Bem, D. J., & Honorton, C. (1994). Does psi exist? Replicable evidence for an anomalous process of information transfer. *Psychological Bulletin*, 115, 4–18.
- Cohen, J. (1994). The Earth is round ($p < .05$). *American Psychologist*, 49, 997–1003. <http://dx.doi.org/10.1037/0003-066X.49.12.997>
- Evanschitzky, H., Baumgarth, C., Hubbard, R., & Armstrong, J. S. (2007). Replication research's disturbing trend. *Journal of Business Research*, 60, 411–415. <http://dx.doi.org/10.1016/j.jbusres.2006.12.003>
- Furchtgott, E. (1984). Replicate, again and again. *American Psychologist*, 39, 1315–1316. <http://dx.doi.org/10.1037/0003-066X.39.11.1315.b>
- Hyman, R. (1994). Anomaly or artifact? Comments on Bem and Honorton. *Psychological Bulletin*, 115, 19–24. <http://dx.doi.org/10.1037/0033-2909.115.1.19>
- Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLOS Medicine*, 2(8), e124. <http://dx.doi.org/10.1371/journal.pmed.0020124>
- Lykken, D. T. (1968). Statistical significance in psychological research. *Psychological Bulletin*, 70, 151–159. <http://dx.doi.org/10.1037/h0026141>
- Makel, M. C., & Plucker, J. A. (2014). Facts are more important than novelty: Replication in the education sciences. *Educational Researcher*, 43, 304–316. <http://dx.doi.org/10.3102/0013189X14545513>

- Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in psychology research: How often do they really occur? *Perspectives on Psychological Science*, 7, 537–542. <http://dx.doi.org/10.1177/1745691612460688>
- Sutton, J. (Ed.). (2012). Replication: Psychology's house of cards? [Special issue]. *The Psychologist*, 25(5).
- Rosenthal, R. (1966). *Experimenter effects in behavioral research*. East Norwalk, CT: Appleton-Century-Crofts.
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86, 638–641. <http://dx.doi.org/10.1037/0033-2909.86.3.638>
- Schmidt, S. (2009). Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Review of General Psychology*, 13, 90–100.
- Wagenmakers, E. J., Wetzels, R., Borsboom, D., & van der Maas, H. (2011). Why psychologists must change the way they analyze their data: The case of psi: Comment on Bem (2011). *Journal of Personality and Social Psychology*, 100, 426–432. <http://dx.doi.org/10.1037/a0022790>