American Psychological Association

Volume 128
Number 6

August 2019

Published eight times

ISSN 0021-843X

Journal of Abnormal Psychology

Founded by Morton Prince, 1906

www.apa.org/pubs/journals/abn

Journal of Abnormal Psychology®

www.apa.org/pubs/journals/abn

August 2019

Volume 128
Number 6

Special Section: Increasing Replicability, Transparency, and Openness in Clinical Psychological Research
Editors: Jennifer L. Tackett and Joshua D. Miller

(*Contents continue*)

Transdiagnostic Research & Epidemeology

(*Contents continue*)

(*Contents continue*)

Disorders of Thought & Mania

Other

American Psychological Association

Journal of Abnormal Psychology

0021-843X/19/$12.00

2019, Vol. 128, No. 6, 487–492
http://dx.doi.org/10.1037/abn0000455

COMMENTARY

Introduction to the Special Section on Increasing Replicability, Transparency, and Openness in Clinical Psychology

Jennifer L. Tackett
Northwestern University

Joshua D. Miller
University of Georgia

As psychological research comes under increasing fire for the crisis of replicability, attention has turned to methods and practices that facilitate (or hinder) a more replicable and veridical body of empirical evidence. These trends have focused on open science initiatives, including an emphasis on replication, transparency, and data sharing. Despite this broader movement in psychology, clinical psychologists and psychiatrists have been largely absent from the broader conversation on documenting the extent of existing problems as well as generating solutions to problematic methods and practices in our area (Tackett et al., 2017). The goal of the current special section was to bring together psychopathology researchers to explore these and related areas as they pertain to the types of research conducted in clinical psychology and allied disciplines.

General Scientific Summary
Current attention to replicability problems in psychological science has uncovered many methods, practices, and cultural infrastructures in need of scientific reform. Although largely uninvolved in this movement to date, the current special section brings clinical psychology and psychopathology researchers to the table to address a number of these concerns.

Keywords: open science, replication, reproducibility, methodology, clinical psychology

The replication crisis that has emerged in psychology, and the social and biomedical sciences more broadly, is reasonably well known by now (e.g., Lilienfeld, 2017; Shrout & Rodgers, 2018; Spellman, 2015). Yet the lack of engagement by the clinical psychology community, already previously noted (e.g., Tackett et al., 2017; Tackett, Brandes, King, & Markon, 2019), remains a glaring and disconcerting omission on the broader replicability landscape. The importance of these issues for clinical psychology and allied disciplines (e.g., psychiatry) with regard to the need for reliable and valid data on the conceptualization, assessment, diagnosis, and treatment of mental health difficulties seems obvious, so the relative lack of engagement by the community is both perplexing and troubling. Of the many insights that have emerged following identification of problems with replicability and current practices, one of the most critical is that business as usual in psychology methods and practices has important consequences for the formation of a robust, replicable scientific literature. Yet the field of clinical psychology has largely proceeded with business as usual in recent years. To continue moving our discipline to a stage of self-examination and scientific progress, this special section offers a number of focused efforts at advancing the current state of our science.

In preparing the call for this special section, we focused broadly on relevant issues, including replicability of findings in our field; transparency of our scientific hypotheses and practices; and openness of our data, measures, and scientific efforts. We invited abstracts for papers on any of these topics, including metascientific papers on these themes as well as hands-on exemplar papers illustrating these practices as applied to data or problems in clinical psychology. The response to the initial call was impressive and gave some cause for optimism, with scores of abstracts submitted for consideration. The articles published in this section represent a sampling of these submissions, with an attempt to achieve breadth across relevant topics. It is our hope that the response to this call reflects more than interest in these topics from the broader com-

Editor's Note. This is an introduction to the special section "Increasing Replicability, Transparency, and Openness in Clinical Psychological Research." Please see the Table of Contents here: https://psycnet.apa.org/journals/abn/128/6.—AM

Jennifer L. Tackett, Department of Psychology, Northwestern University; Joshua D. Miller, Department of Psychology, University of Georgia.

Correspondence concerning this article should be addressed to Jennifer L. Tackett, Department of Psychology, Northwestern University, 2029 Sheridan Road, Evanston, IL 60208. E-mail: jenniferltackett@gmail.com or jennifer.tackett@northwestern.edu

American Psychological Association

Journal of Abnormal Psychology

0021-843X/19/$12.00

2019, Vol. 128, No. 6, 493–499
http://dx.doi.org/10.1037/abn0000435

An *N*-Pact Factor for Clinical Psychological Research

Kathleen W. Reardon, Avanté J. Smack, Kathrin Herzhoff, and Jennifer L. Tackett
Northwestern University

Although an emphasis on adequate sample size and statistical power has a long history in clinical psychological science (Cohen, 1992), increased attention to the replicability of scientific findings has renewed focus on the importance of statistical power (Bakker, van Dijk, & Wicherts, 2012). These recent efforts have not yet circled back to modern clinical psychological research, despite the importance of sample size and power in producing a credible body of evidence. As one step in this process of scientific self-examination, the present study estimated an *N*-pact Factor (the statistical power of published empirical studies to detect typical effect sizes; Fraley & Vazire, 2014) in 2 leading clinical journals (the *Journal of Abnormal Psychology* [*JAP*] and *the Journal of Consulting and Clinical Psychology* [*JCCP*]) for the years 2000, 2005, 2010, and 2015. Study sample size, as one proxy for statistical power, is a useful focus because it allows comparisons with other subfields and may highlight some of the core methodological differences between clinical and other areas. We found that, across all years examined, the average median sample size in clinical research was 179 participants (175 for *JAP* and 182 for *JCCP*). The power to detect a small to medium effect size of .20 is just below 80% for both journals. Although the clinical *N*-pact factor was higher than that estimated for social psychology, the statistical power in clinical journals is still limited to detect many effects of interest to clinical psychologists, with little evidence of improvement in sample sizes over time.

General Scientific Summary
Statistical power is a key indicator of whether a body of research is trustworthy. This study evaluated the median sample size of published research in 2 clinical psychology journals. Overall, the average median sample size for clinical psychology was greater than that calculated for social psychology, but the power in clinical journals is still limited.

Keywords: power, sample size, replicability, clinical psychology, metascience

Over the past several years, psychological science has begun to engage in serious self-examination, with the aim of creating a more rigorous and trustworthy body of evidence. Although the dangers of inadequate statistical power have been discussed by psychologists for decades (Cohen, 1962; Rossi, 1990), this self-examination has led to a renewed focus on bolstering study power through larger samples as a method of protecting against false positive findings (e.g., Ioannidis, 2005; Szucs & Ioannidis, 2017). Typical power in social and personality psychology has been described by Fraley and Vazire's (2014) *N*-pact Factor (NF), a measure of the median sample size of published empirical studies. Typical statistical power in clinical psychology, however, has not yet been examined. In fact, clinical psychology has largely been left out of the ongoing conversation about improving the science (see Tackett et al., 2017) and faces some unique obstacles directly relevant to calls for increased sample size. Thus, this article represents an extension of Fraley and Vazire's NF approach to the clinical psychological literature.

Power is defined as the likelihood that a study can detect a true effect if one exists (Cohen, 1992). Studies are ideally designed to have 80% power to detect an effect, although in practice, they often have lower power because of smaller sample sizes or a misunderstanding of how statistical significance testing and power are interdependent (Greenland et al., 2016). Work in other subfields has highlighted a glut of potentially false-positive findings (e.g., Simmons, Nelson, & Simonsohn, 2011). Although never desirable, we argue that false-positive findings in clinical psychology may lead to even more wasted time, effort, and resources. Adequate power is one of the best ways to minimize the risk for false positives (Marino, 2018). Sample size (*N*), although only one

Editor's Note. Angus MacDonald III served as the action editor for this article.—AM

Kathleen W. Reardon, Avanté J. Smack, Kathrin Herzhoff, and Jennifer L. Tackett, Department of Psychology, Northwestern University.

A summary of this work was presented as a talk by Jennifer L. Tackett at the Society for Research in Psychopathology annual meeting in September 2018. A preprint version of this article is available on PsyArXiv (psyarxiv.com/4fybk).

Correspondence concerning this article should be addressed to Kathleen W. Reardon or Jennifer L. Tackett, Department of Psychology, Northwestern University, Swift Hall 102, 2029 Sheridan Road, Evanston, IL 60208. E-mail: kwreardon@u.northwestern.edu or jennifer.tackett@northwestern.edu

American Psychological Association

Journal of Abnormal Psychology

0021-843X/19/$12.00

2019, Vol. 128, No. 6, 500–509
http://dx.doi.org/10.1037/abn0000421

Evaluating the Evidential Value of Empirically Supported Psychological Treatments (ESTs): A Meta-Scientific Review

John Kitchener Sakaluk
University of Victoria

Alexander J. Williams
University of Kansas

Robyn E. Kilshaw
University of Utah

Kathleen Teresa Rhyner
Canandaigua VA Medical Center, Canandaigua, New York

Empirically supported treatments (or therapies; ESTs) are the gold standard in therapeutic interventions for psychopathology. Based on a set of methodological and statistical criteria, the APA has assigned particular treatment-diagnosis combinations EST status and has further rated their empirical support as *Strong*, *Modest*, and/or *Controversial*. Emerging concerns about the replicability of research findings in clinical psychology highlight the need to critically examine the evidential value of EST research. We therefore conducted a metascientific review of the EST literature, using clinical trials reported in an existing online APA database of ESTs, and a set of novel evidential value metrics (i.e., rates of misreported statistics, statistical power, R-Index, and Bayes Factors). Our analyses indicated that power and replicability estimates were concerningly low across almost all ESTs, and individually, some ESTs scored poorly across multiple metrics, with *Strong* ESTs failing to continuously outperform their *Modest* counterparts. Lastly, we found evidence of improvements over time in statistical power within the EST literature, but not for the strength of evidence of EST efficacy. We describe the implications of our findings for practicing psychotherapists and offer recommendations for improving the evidential value of EST research moving forward.

General Scientific Summary
This review suggests that although the underlying evidence for a small number of empirically supported therapies is consistently strong across a range of metrics, the evidence is mixed or consistently weak for many, including some classified by Division 12 of the APA as "Strong."

Keywords: empirically supported treatments, evidential value, metascience, replicability

Clinical efficacy underpins everything the psychotherapy industry promises and is ethically bound to deliver. Questions about whether psychotherapy works—and if so, for whom and under what conditions—have guided research for the better part of a century (e.g., Eysenck, 1952). However, only since the 1970s have controlled trials like those used in medicine flourished in the field of psychology (e.g., Klerman, Dimascio, Weissman, Prusoff, & Paykel, 1974). In an effort to better synthesize and disseminate the results of this efficacy research, an APA Division 12 Task Force on Promotion and Dissemination of Psychological Procedures (1995) created a continually updated list of the therapies that have reached a certain level of research support for treating patients with specific diagnoses. These therapies came to be known as *empirically supported psychological treatments* (or *therapies*) (ESTs; Kendall, 1998).

John Kitchener Sakaluk, Department of Psychology, University of Victoria; Alexander J. Williams, Department of Psychology, University of Kansas; Robyn E. Kilshaw, Department of Psychology, University of Utah; Kathleen Teresa Rhyner, Canandaigua VA Medical Center, Canandaigua, New York.

Alexander J. Williams and John Kitchener Sakaluk share first authorship; their order was arbitrarily decided by a bitterly contested coin toss. The authors sincerely thank Drs. Corker, Fried, Grubbs, Kirk, and Nuijten for their consultation on methodological and clinical matters throughout our review, and Kyle Dirck and Robert Williams for their helpful suggestions on earlier drafts of the manuscript.

This research was supported by a SSHRC Insight Development Grant awarded to John Kitchener Sakaluk. He is a statistical consultant at other journals and methodological/quantitative expert who has published on the replicability crisis before. Alexander J. Williams is a psychological clinic director who has used ESTs in his own practice and encourages trainees to use them. Kathleen Teresa Rhyner is a VA psychologist.

Correspondence concerning this article should be addressed to John Kitchener Sakaluk, Department of Psychology, University of Victoria, P. O. Box 1700 STN CSC, Victoria, British Columbia, Canada V8W 2Y2 or to Alexander J. Williams, Department of Psychology, University of Kansas, 1415 Jayhawk Boulevard, Fraser Hall, Room 426, Lawrence, KS 66045-7556. E-mail: jksakaluk@gmail.com or alexwilliams@ku.edu

American Psychological Association

Journal of Abnormal Psychology

© 2019 American Psychological Association
0021-843X/19/$12.00

2019, Vol. 128, No. 6, 517–527
http://dx.doi.org/10.1037/abn0000424

A Step-By-Step Guide on Preregistration and Effective Data Sharing for Psychopathology Research

Angelos-Miltiadis Krypotos, Irene Klugkist, Gaëtan Mertens, and Iris M. Engelhard
Utrecht University

Data analysis in psychopathology research typically entails multiple stages of data preprocessing (e.g., coding of physiological measures), statistical decisions (e.g., inclusion of covariates), and reporting (e.g., selecting which variables best answer the research questions). The complexity and lack of transparency of these procedures have resulted in two troubling trends: the central hypotheses and analytical approaches are often selected after observing the data, and the research data are often not properly indexed. These practices are particularly problematic for (experimental) psychopathology research because the data are often hard to gather due to the target populations (e.g., individuals with mental disorders), and because the standard methodological approaches are challenging and time consuming (e.g., longitudinal studies). Here, we present a workflow that covers study preregistration, data anonymization, and the easy sharing of data and experimental material with the rest of the research community. This workflow is tailored to both original studies and secondary statistical analyses of archival data sets. In order to facilitate the implementation of the described workflow, we have developed a free and open-source software program. We argue that this workflow will result in more transparent and easily shareable psychopathology research, eventually increasing and replicability reproducibility in our research field.

General Scientific Summary
We describe a workflow for preregistering as well as for sharing data and materials of psychopathology studies. To facilitate the implementation of this workflow, we also present a free, easy to use, software we have recently developed.

Keywords: replicability, reproducibility, experimental psychopathology, R

The main goals of psychopathology research are to unveil the factors that contribute to the genesis and maintenance of mental disorders, and to develop relevant prevention and intervention programs (Marks & Yardley, 2004; Van den Hout, Engelhard, & McNally, 2017). This research area often requires challenging data accumulation methods (Comer & Kendall, 2013), including longitudinal research in samples at risk of developing mental disorders, and demanding research protocols (e.g., randomized control trials [RCTs]). Given these challenges, it is crucial to make the most of the collected data.

The timely answering of research questions depends on how reliable the published literature is. Recent findings in psychology, however, suggest that many popular effects cannot be reproduced (e.g., Open Science Collaboration, 2015; Świątkowski & Dompnier, 2017). There are scientific, ethical, and practical reasons that make such low reproducibility deleterious for psychopathology research. Scientifically, a finding with low reproducibility is not informative, and it slows the progress of our field. Ethically, unreliable research findings stall the development of effective interventions for mental disorders. Practically, unreproducible psychopathology research is a waste of resources and patients' time (Baker, McFall, & Shoham, 2008). Arguably, psychopathology research can only progress by studies that are replicable (i.e., repetition of the results using similar procedures but a new data set) and reproducible (i.e., obtaining the same results as the original study by using the same procedures and data; Brandt et al., 2014; Goodman, Fanelli, & Ioannidis, 2016).

Replicability in psychology is often hampered by the formation of a study's hypotheses after the results are already known (Kerr, 1998; Nosek, Spies, & Motyl, 2012). Hypothesizing based on the results could inflate the rate of false positives and lead to nonin-

Angelos-Miltiadis Krypotos, Department of Clinical Psychology, Utrecht University; Irene Klugkist, Department of Methodology and Statistics, Utrecht University; Gaëtan Mertens and Iris M. Engelhard, Department of Clinical Psychology, Utrecht University.

Angelos-Miltiadis Krypotos is currently also affiliated with Department of Health Psychology, KU Leuven.

This research was supported by a Vici grant awarded by the Netherlands Organization for Scientific Research (NWO; Grant 453-15-005) to IME. The authors thank Nicolas Perez for programming, together with Angelos-Miltiadis Krypotos, for the software presented in this paper.

Correspondence concerning this article should be addressed to Angelos-Miltiadis Krypotos, Department of Clinical Psychology, Utrecht University, Heidelberglaan 1, 3584 CS Utrecht, the Netherlands. E-mail: amkrypotos@gmail.com

American Psychological Association

Journal of Abnormal Psychology

0021-843X/19/$12.00

2019, Vol. 128, No. 6, 528–540
http://dx.doi.org/10.1037/abn0000451

The Registration Continuum in Clinical Science: A Guide Toward Transparent Practices

Stephen D. Benning
University of Nevada, Las Vegas

Rachel L. Bachrach
VA Pittsburgh Healthcare System, Pittsburgh, Pennsylvania

Edward A. Smith and Andrew J. Freeman
University of Nevada, Las Vegas

Aidan G. C. Wright
University of Pittsburgh

Clinical scientists can use a continuum of registration efforts that vary in their disclosure and timing relative to data collection and analysis. Broadly speaking, registration benefits investigators by offering stronger, more powerful tests of theory with particular methods in tandem with better control of long-run false positive error rates. Registration helps clinical researchers in thinking through tensions between bandwidth and fidelity that surround recruiting participants, defining clinical phenotypes, handling comorbidity, treating missing data, and analyzing rich and complex data. In particular, registration helps record and justify the reasons behind specific study design decisions, though it also provides the opportunity to register entire decision trees with specific endpoints. Creating ever more faithful registrations and standard operating procedures may offer alternative methods of judging a clinical investigator's scientific skill and eminence because study registration increases the transparency of clinical researchers' work.

General Scientific Summary
Study registration allows clinical scientists to make their work more credible and transparent to fellow researchers and the general public. We describe the dimensions of disclosure and timing relative to data collection and analysis on which researchers can register their study designs, how registration makes for better science, and the kinds of issues for which registration is particularly helpful in clinical research. We also show how registration permits researchers to flexibly register specific individual decisions or complex decision trees to deal with potential problems in a study.

Keywords: preregistration, coregistration, postregistration, transparency, flexibility

Registration is the act of formally recording the components of a scientific study (e.g., aims, hypotheses, methods, data analytic strategy) in an official capacity (e.g., to a journal for peer review or a public repository). The goals of registration include increasing research's transparency (Gernsbacher, 2018) and maintaining the diagnostic value and credibility of a study's confirmatory analyses (Wagenmakers, Wetzels, Borsboom, van der Maas, & Kievit, 2012). Registration has gained steam in the last half decade following concerns regarding research practices that encouraged blurring the lines between prediction and postdiction (Nosek, Ebersole, DeHaven, & Mellor, 2018). This scientific ecology has led to an unfortunate plethora of studies whose findings could not be replicated in psychology (Open Science Collaboration, 2015) and other clinical disciplines (e.g., Begley & Ellis, 2012). Registration poses special challenges for clinical science (Tackett et al., 2017), a discipline that is uniquely poised to benefit people's lives and ameliorate suffering directly through research.

Registration is applicable to a broad range of clinical science from bench to bedside to help produce replicable work. Furthermore, not all registrations of scientific work must—or can—be accomplished before the first participant is run. We describe a continuum of registration approaches clinical scientists can use, from preregistered plans filed in anticipation of commencing a study to postregistrations of analyses involving complex archival

Stephen D. Benning, Department of Psychology, University of Nevada, Las Vegas; Rachel L. Bachrach, Interdisciplinary Addiction Program for Education and Research, Center for Health Equity and Research Promotion, Mental Illness Research, Education and Clinical Center, VA Pittsburgh Healthcare System, Pittsburgh, Pennsylvania; Edward A. Smith and Andrew J. Freeman, Department of Psychology, University of Nevada, Las Vegas; Aidan G. C. Wright, Department of Psychology, University of Pittsburgh.

Aidan G. C. Wright's effort on this project was supported by the National Institute of Mental Health (L30 MH101760). The opinions expressed are solely those of the authors and not those of the funding source.

Correspondence concerning this article should be addressed to Stephen D. Benning, Department of Psychology, University of Nevada, Las Vegas, 4505 South Maryland Parkway, Box 455030, Las Vegas, NV 89154-5030, or to Aidan G. C. Wright, Department of Psychology, University of Pittsburgh, 4119 Sennott Square, 210 South Bouquet Street, Pittsburgh, PA 15260. E-mail: stephen.benning@unlv.edu or aidan@pitt.edu

American Psychological Association

Journal of Abnormal Psychology

0021-843X/19/$12.00

2019, Vol. 128, No. 6, 541–546
http://dx.doi.org/10.1037/abn0000426

Should Psychologists Sign Their Reviews? Some Thoughts and Some Data

Donald R. Lynam
Purdue University

Courtland S. Hyatt
University of Georgia

Christopher J. Hopwood
University of California, Davis

Aidan G. C. Wright
University of Pittsburgh

Joshua D. Miller
University of Georgia

The Open Science Movement (OSM) emphasizes increased transparency at many of the steps in the scientific process and has improved psychological science. In the present article, we discuss whether such transparency should find its way into the review process. We discuss a priori thoughts and intuitions about the costs and benefits of signing reviews. In terms of benefits, these include greater alignment with OSM and greater accountability leading to increases in civility, care, and thoughtfulness of reviews. The most obvious cost is potential retaliation for negative reviews. To check these intuitions, we surveyed a sample of 358 faculty members about their experience and views on signing reviews. Results both underscored and extended the initial intuitions. Results suggest there are many benefits to increasing the incidence of reviewers signing their reviews. Fears of retaliation seem to be somewhat exaggerated. We discuss possible means of reducing the possibility of retaliation.

General Scientific Summary
This study suggests that most people in the field view signing reviews as a part of the open science framework. Surveyed faculty members revealed a number of benefits and potential costs to signing. There are several steps that might be taken to reduce the costs.

Keywords: open science, peer review, transparency, signing reviews

The Open Science Movement (OSM) has gained popularity as researchers come to understand the power it has to improve our science(s), amid growing concerns regarding the replicability of health sciences (e.g., Elbaz et al., 2006), psychology in general (Open Science Collaboration, 2015), and clinical psychology more specifically (Hopwood & Vazire, in press; Tackett et al., 2017). OSM adherents have pushed for increased transparency at all steps in the scientific process, including information as to how sample sizes were determined, methods employed, analytic choices, and increased access to data after they have been collected. Practices such as preregistration, registered reports, and the sharing of data and analytic code render researchers accountable to the field and themselves. In addition to increasing accountability, the OSM may also increase thoughtfulness. To preregister a study, psychologists must think through issues such as sample size, measurement, effect of interest, statistical approach, and procedures for handling invalid data in advance.

Discussions about transparency in the review process have not paralleled discussions about transparency in the research process. The double-blind review process does not align well with the emphasis on transparency and accountability in the OSM. Roediger (2018) recently considered the issue, noting, "Transparent practices seem here to stay. . . . With one glaring exception: Transparency in publication practices." In his article, he touched briefly on the issues involved in the anonymity of reviews, action letters, and submissions. In this article, we discuss in more detail the cases for and against signing reviews and provide some initial survey data as to how research psychologists feel about signing reviews and its place in the OSM.

Editor's Note. Angus MacDonald III served as the action editor for this article.—AM

Donald R. Lynam, Department of Psychological Sciences, Purdue University; Courtland S. Hyatt, Department of Psychology, University of Georgia; Christopher J. Hopwood, Department of Psychology, University of California, Davis; Aidan G. C. Wright, Department of Psychology, University of Pittsburgh; Joshua D. Miller, Department of Psychology, University of Georgia.

Study procedures were approved by the University of Georgia IRB (Protocol 00005836: Transparency in the Peer Review Process).

Correspondence concerning this article should be addressed to Donald R. Lynam, Department of Psychological Sciences, Purdue University, 703 Third Street, West Lafayette, IN 47907-2081. E-mail: dlynam@purdue.edu

American Psychological Association

Journal of Abnormal Psychology

0021-843X/19/$12.00

2019, Vol. 128, No. 6, 547–562
http://dx.doi.org/10.1037/abn0000417

Using Implementation Science to Close the Gap Between the Optimal and Typical Practice of Quantitative Methods in Clinical Science

Kevin M. King, Michael D. Pullmann, Aaron R. Lyon, and Shannon Dorsey
University of Washington

Cara C. Lewis
Kaiser Permanente Washington Health Research Institute, Seattle, Washington

Quantitative methods remain the fundamental approach for hypothesis testing, but in approaches to data analysis there is substantial evidence of a gap between what is optimal and what is typical. It is clear that diffusion and dissemination alone are not maximally effective at improving data analytic practices in clinical psychological science. Amid declines in quantitative psychology training, and growing demand for advanced quantitative methods, applied researchers are increasingly called upon to conduct and evaluate research using methods in which they lack expertise. This "research-to-practice" gap in which rigorously developed and empirically supported quantitative methods are not applied in practice has received little attention. In this article, we describe how implementation science, which aims to reduce the research-to-practice gap in health care, offers a promising set of methods for closing the gap for quantitative methods. By identifying determinants of practice (i.e., barriers and facilitators of change), implementation strategies can be selected to increase adoption and high-fidelity application of new quantitative methods to improve scientific inferences and policy and practice decisions in clinical psychological science.

General Scientific Summary

Making studies more replicable will require more effective use of statistics in research, but there is a large gap between how statistics are applied in psychological research and how they should be applied. The current article describes how the lessons of clinical implementation science, which has focused on getting evidence based treatments into community practice settings, may be applied to improve research in clinical psychology.

Keywords: quantitative methods, open science, implementation science, quantitative implementation

There is a gap between how psychological science might be optimally conducted and how it is typically conducted, which undermines the credibility of research findings. For example, although 97% of effects in a large scale replication effort were statistically significant in their original studies, only 36% of those effects were statistically significant in new, larger samples, with a median effect size half that of the original studies (Aarts et al., 2015), echoing earlier concerns that many published research findings may be false (Ioannidis, 2005). There are many factors that likely undermine the credibility of current research, but many of them reflect a gap between how research should be conducted and how it is conducted. Researchers engage in questionable research practices (Fiedler & Schwarz, 2016; John, Loewenstein, & Prelec, 2012) to achieve statistically significant results through selective testing, reporting, or utilizing other "researcher degrees of freedom" which can dramatically increase the number of false positive findings (Simmons, Nelson, & Simonsohn, 2011). This leads to higher than expected rates of false positives, especially when research practices are not robust (Ioannidis, 2005; Szucs, 2016). Compounding the problems, researchers then engage in hypothesizing after the results are known (i.e., HARKing) implying that post hoc findings are actually a priori (Kerr, 1998).

Clinical psychological science has long faced a similar gap between what researchers operationalize as optimal treatment and what occurs in community practice settings (Kazdin, 2008). Researchers have long known that psychotherapy has large effect sizes when conducted in controlled research settings (Weisz, Donenberg, Han, & Weiss, 1995), but fails to have much if any effectiveness in typical clinical samples and settings (Weisz & Jensen, 2001). Indeed, many fields have struggled to translate research identifying best practices into meaningful change in real-world outcomes (Damschroder et al., 2009). For example, studies have reported that it can take 10 to 25 years for biomedical

Kevin M. King, Michael D. Pullmann, Aaron R. Lyon, and Shannon Dorsey, Department of Psychology, University of Washington; Cara C. Lewis, Kaiser Permanente Washington Health Research Institute, Seattle, Washington.

We appreciate the feedback from Teague Henry and Dale Kim who provided comments on an earlier version of this article.

Correspondence concerning this article should be addressed to Kevin M. King, Department of Psychology, University of Washington, Box 351525, Seattle, WA 98195-1525. E-mail: kingkm@uw.edu

American Psychological Association

Journal of Abnormal Psychology

0021-843X/19/$12.00

2019, Vol. 128, No. 6, 563–573
http://dx.doi.org/10.1037/abn0000406

A Meta-Analysis of Bias at Baseline in RCTs of Attention Bias Modification: No Evidence for Dot-Probe Bias Towards Threat in Clinical Anxiety and PTSD

Anne-Wil Kruijt
Stockholm University

Sam Parsons and Elaine Fox
University of Oxford

Considerable effort and funding have been spent on developing Attention Bias Modification (ABM) as a treatment for anxiety disorders, theorized to exert therapeutic effects through reduction of a tendency to orient attention toward threat. However, meta-analytical evidence that clinical anxiety is characterized by threat-related attention bias is thin. The largest meta-analysis to date included dot-probe data for $n = 337$ clinically anxious individuals. Baseline measures of biased attention obtained in ABM RCTs form an additional body of data that has not previously been meta-analyzed. This article presents a meta-analysis of threat-related dot-probe bias measured at baseline for 1,005 clinically anxious individuals enrolled in 13 ABM RCTs. Random-effects meta-analysis indicated no evidence that the mean bias index (BI) differed from zero ($k = 13$, $n = 1005$, mean BI = 1.8 ms, $SE = 1.26$ ms, $p = .144$, 95% confidence interval [−0.6, 4.3]. Additional Bayes factor analyses also supported the point-zero hypothesis (BF10 = .23), whereas interval-based analysis indicated that mean bias in clinical anxiety is unlikely to extend beyond the 0 to 5 ms interval. Findings are discussed with respect to strengths (relatively large samples, possible bypassing of publication bias), limitations (lack of control comparison, repurposing data, specificity to dot-probe data), and theoretical and practical context. We suggest that it should no longer be assumed that clinically anxious individuals are characterized by selective attention toward threat. Clinically anxious individuals enrolled in RCTs for Attention Bias Modification are not characterized by threat-related attention bias at baseline.

General Scientific Summary

It is widely believed that anxiety is characterized by a tendency to orient attention specifically toward threatening information and that this tendency (called attention bias) can be measured using a computer task called the "dot-probe task." Over the past decade, studies have tested whether a training version of this task can be used to modify bias, which might then be used as a new treatment (Attention Bias Modification). We analyzed levels of attention bias measured before participants started the modification training in 13 studies enrolling 1,005 diagnosed anxious patients. We found no evidence that clinically anxious people are characterized by attention bias toward threat.

Keywords: attention bias, clinical anxiety, meta-analysis, attention bias modification, translational research

Supplemental materials: http://dx.doi.org/10.1037/abn0000406.supp

Anne-Wil Kruijt, Department of Psychology, Stockholm University; Sam Parsons and Elaine Fox, Department of Experimental Psychology, University of Oxford.

This article describes a meta-analytical study. Because all data originates from previous clinical trials, this study was deemed exempt from ethics committee approval.

The research reported here was funded by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007–2013)/ERC grant agreement [324176]. Preliminary analyses on a smaller subset of studies were presented at the 9th Swedish Congress on Internet interventions (SWEsrii) in November 2017. A preprint version of this article is available at https://psyarxiv.com/rfjup.

Correspondence concerning this article should be addressed to Anne-Wil Kruijt, Department of Psychology, Stockholm University, SE-106 91 Stockholm, Sweden. E-mail: anne-wil.kruijt@psychology.su.se or mail@awkruijt.nl