C

2016 Annual Convention - Symposium/Other Proposal

Submitter: Dennis Wendt

sym16108: Beyond the "Replication Crisis": Diverse Considerations for Psychology's Future

Type of program:	Symposia
Title of program:	Beyond the "Replication Crisis": Diverse Considerations for Psychology's Future
First index term:	50 Methods and Measurement
Second index term:	58 Philosophical/Theoretical
Theme:	
Brief Content Description:	Psychology's recent "replication crisis" is an opportunity to critically address and creatively improve the future of psychological research. This symposium includes diverse presentations and interaction among theoretical and methodological experts.
Division to submit:	24 - Theoretical and Philosophical
Other division(s) appropriate for submission:	01 - Society for General Psychology; 05 - Quantitative and Qualitative Methods; 08 - Society for Personality and Social Psychology; 26 - Society for the History of Psychology; 45 - Psychological Study of Culture, Ethnicity and Race; 52 - International Psychology
Theme:	The Future of Psychology: Advancing the Field in a Rapidly Changing World
Length of program:	1 hr. 50 min.
General statement:	In a highly visible 2015 "Science" article, the Open Science Collaboration (OSC) demonstrated that experiments from only 34% of original studies published in top-tier psychology journals could be replicated with statistically significant results. This discovery—considered by some to be a "crisis"—suggests a challenge to the reliability of psychological research and the discipline's ambition to discover generalizable findings across populations. At the same time, this revelation presents an opportunity—at a ripe time, in light of the Hoffman torture report—to humbly reflect upon and creatively refine and re-imagine psychology's methodological and scientific production practices. Much is at stake for the discipline and those it serves, and thus it behooves us to consult diverse reflections from multiple perspectives. To this end, this symposium is carefully designed to maximize the diversity of perspectives that are given voice as well as to stimulate discussion and audience participation. Psychologists representing a diversity of theoretical, methodological, and sub-disciplinary perspectives (including two OSC members and four presenters from outside the U.S.) will give brief presentations (5-6 minutes each), in which they focus on recommendations for future psychological research. These consist of an overview of the wide-ranging responses to the "Science" article along with major implications/recommendations (Cody Christopherson [OSC]); presentations on technological and methodological recommendations for improving publication practices and methodological rigor (Etienne LeBel [OSC], Scott Hofer, Brent Slife); presentations addressing challenges with particular disciplinary domains, namely clinical science (Scott Lilienfeld), qualitative inquiry (Ruthellen Josselson), and international/cultural contexts (Steven Heine, Louise Sundararajan); and presentations that address "big picture" criticisms and challenges for the future of psychology (Brian Schiff, Lisa Osbeck). The remaining time will consist of moderated discussion among presenters (25-30 minutes) and responses to audience questions (20 minutes), allowing for critical examination of the ideas presented.
Accommodation/Setup	To facilitate conversation with a large number of presenters, we would prefer a setup that would allow for 10 presenters to sit facing the audience, along with multiple

request:	microphones (if possible).
Submit For CE:	Yes
Learning Objectives:	explain at least three major implications of the recently demonstrated low rate of replication among psychological studies. describe at least five major recommendations for addressing the low rate of replication among psychological studies.
Submitter:	Brian Schiff (Chair)

Chair 1:	Dennis C Wendt, PhD *University of Washington School of Medicine, Seattle, WA*
E-Mail address:	dcwendt@uw.edu
Mailing address:	Department of Psychiatry and Behavioral Sciences, University of Washington School of Medicine, 1107 NE 45th St., Suite 120, Box 354805, Seattle, WA 98105-4631
Phone numbers:	206-616-6131 (office), 801-636-5823 (home), 801-636-5823 (cell)
Membership status:	APA Member

Chair 2:	Brian Schiff, PhD *American University of Paris, Paris, France*
E-Mail address:	bschiff@aup.edu
Mailing address:	Department of Psychology, American University of Paris, 5 Boulevard de la Tour-Maubourg, Paris, France
Membership status:	APA Member

Participant 1:	Cody D Christopherson, PhD *Southern Oregon University, Ashland, OR*
E-Mail address:	christoc1@sou.edu
Mailing address:	Psychology Department, Southern Oregon University, 1250 Siskiyou Blvd., Ashland, OR 97520
Title of presentation:	Never Let a Replication Crisis Go to Waste: Improving Psychological Science Practices with Science
Summary:	The article "Estimating the Reproducibility of Psychological Science" has caused a ripple effect of responses from psychology researchers, scientists outside of psychology, and the general public. According to one tracker of research media attention, this article has had the seventh most media impact of all tracked articles ever published by Science. These responses range from declaring psychology dead to denying that anything surprising or troubling has been found. In this presentation, I will briefly outline the range of responses to the project and provide background about deliberate choices made by Open Science Collaboration that helps make sense of this range of responses. I will also explain what I see as a key implication of the findings. A greater self-awareness of assumptions embedded in our research methods and publishing practices as well as promotion & tenure incentives should lead to substantial changes in these processes. Psychological science has been narrowly defined and meta-investigation has been prematurely restricted (or made unnecessarily difficult) with

	these implicit assumptions, to the detriment of the field. Change is possible but will require a grassroots shift in the gestalt, including the media, researchers, publishers, consumers of research, and those who fund and implement research, and not an exclusive reliance on those who have been successful in the current system.
Electronic archiving:	Yes
Membership status:	APA Member

Participant 2:	Etienne P LeBel, PhD *University of Western Ontario / Open Science Collaboration, London, ON, Canada*
E-Mail address:	etienne.lebel@gmail.com
Mailing address:	Department of Psychology, The University of Western Ontario, Social Science Centre, London, ON N6A 3K7 Canada
Title of presentation:	Pre- and Post-Publication Verification of Empirical Results in Psychology: A New Way Forward
Summary:	Science is the most successful system of generating cumulative knowledge about how our world works. The key activity responsible for science's success is independent verification, which involves three facets: (1) peer-evaluation, (2) verification of analyses / analytic reproducibility, and (3) independent replication. These three layers of verification maximize the likelihood of detecting errors, hence maximizing the reliability and validity of empirical results, generating cumulative objective knowledge. The current academic incentive structure, however, does not reward verification and so verification rarely occurs and when it does is highly difficult and inefficient. Curate Science is a web application that facilitates the pre- and post-publication curation and verification of empirical results, as to increase the growth of cumulative knowledge and theoretical progress. The site facilitates verification in terms of verifying the (1) reproducibility of analyses/results and (2) replicability of findings in independent samples. Reproducibility is facilitated by enabling researchers to check and endorse the analytic reproducibility of each other's empirical results via data analyses executed within their web browser. Replicability is facilitated by enabling users to link replications to their original studies with corresponding automatic real-time updating of meta-analytic effect size estimates. The platform also features "revised community abstracts" (crowd-sourced abstracts summarizing how follow-up research has qualified original findings) and curation of organic and external post-publication peer-review commentaries. Curate Science's vision for the future of psychological research is one where verification is routinely and easily done in the cloud, and in which appropriate professional credit is given to researchers who engage in such verification activities (i.e., post-publication peer review and verifying reproducibility and replicability of empirical results).
Electronic archiving:	Yes
Membership status:	Non-Member

Participant 3:	Scott M Hofer, PhD *University of Victoria, Victoria, BC, Canada*
E-Mail address:	smhofer@uvic.ca
Mailing address:	Department of Psychology, University of Victoria, PO Box 1700 STN CSC,

	Victoria, BC V8W 2Y2 Canada
Title of presentation:	International Multiple-Study Research on Aging and Dementia: Lessons Learned
Summary:	The analysis of longitudinal observational data can take many forms and requires many decisions, with research findings and conclusions often found to differ across independent longitudinal studies addressing the same question. Sensitivity of results to minor differences in model, potential for confirmation bias, and overfitting of model to data weaken opportunities for replication research. Differences in measurements, sample composition (e.g., age, cohort, country/culture), and statistical models (e.g., change/time function, covariate set, centering, treatment of incomplete data) also affect the replicability of results. The central aim of the Integrative Analysis of Longitudinal Studies of Aging (IALSA) research network (NIH P01AG043362), comprised of over 100 studies, is to optimize opportunities for replication and cross-validation across heterogeneous sources of longitudinal data by evaluating comparable conceptual and statistical models at the construct-level. Item-level metadata and publicly available analysis scripts enable reproducibility, cross-study replication, and efficient extension of results to alternative models. I reflect on lessons learned from research activities of the IALSA network for the future of replication in psychology with particular attention to large and complex data sets.
Electronic archiving:	Yes
Membership status:	APA Fellow

Participant 4:	Brent D Slife, PhD *Brigham Young University, Provo, UT*
E-Mail address:	slife@byu.edu
Mailing address:	1409 Renaissance Place, Pleasant Grove, UT 84062
Title of presentation:	Problematic Operationalizations: A Piece of the Replication Puzzle
Summary:	As the Open Science Collaboration has demonstrated, psychology has a problem with the replicability of its research results. Roughly two-thirds of even our best research appear not to be replicated. I've written about some philosophy of science reasons for this problem elsewhere (Slife, 2010; 2012; 2014), but the chances of psychological researchers engaging in a radically new logic of method seems remote. Consequently, this presentation focuses on a relatively unexamined approach to raising the number of replications with current quantitative methodologies, specifically increasing the accuracy with which operationalizations represent their constructs. The presentation will first note how virtually all method texts contend that accurate operationalization is pivotal to valid and reliable research (e.g., Krathwohl, 2009; Privatera, 2014). It will then show how rarely research training provides guidance and criteria for accurately formulating and using operationalizations in research. Such training is especially challenging because the relationship between the construct and its operationalization cannot be measured or observed. One part of this relationship, the construct, is not measurable in principle (because it requires operationalization), so the construct's relation to its operationalization cannot itself be observed or measured.

	The upshot is that what is actually studied, the operationalization, may have little relation to the construct intended to be studied, making replication more difficult. The presentation will show how covarying multiple operationalizations can be helpful, but only if certain investigative conditions are met. The presentation concludes by describing a research training regimen that facilitates a more rigorous and accurate formulation of operationalizations, thus allowing more replications.
Electronic archiving:	Yes
Membership status:	APA Fellow

Participant 5:	Scott O Lilienfeld, PhD *Emory University, Atlanta, GA*
E-Mail address:	slilien@emory.edu
Mailing address:	Department of Psychology, Emory University, 36 Eagle Row, Psychology and Interdisciplinary Sciences Building, Room 473, Atlanta, GA 30322
Title of presentation:	Replicability and Clinical Science
Summary:	Over the past decade, growing doubts have emerged concerning the replicability of psychological findings. In particular, a growing cadre of scholars have raised concerns that a large proportion of psychological results are false-positives. Most efforts to ascertain the magnitude of the "replicability problem" in psychology have until recently focused on social and cognitive psychology, with relatively little attention accorded to clinical psychology, psychiatry, and allied disciplines. This omission is unfortunate given that ineffective psychological or psychopharmacological interventions that are erroneously deemed to be efficacious on the basis of unreplicated studies have the potential to do harm, both indirect (opportunity costs) and direct (iatrogenic effects). Moreover, given that many studies in clinical science are extremely time- and labor-intensive owing to difficulties in recruiting difficult-to-access patient populations, the incentives to engage in p-hacking, "harking" (hypothesizing after results are known), and other questionable research practices may be higher in clinical psychology than in most other psychological disciplines. In addition, because many clinical studies require investigators to accrue individuals with low base-rate conditions who fulfill multiple exclusion criteria, statistical power is often a serious concern. Many investigators appear to be unaware that low power not only boosts the likelihood of Type II error, but also increases the odds of false positives owing to the "winner's curse." The increasing influence of the Research Domain Criteria (RDoC) over clinical research may exacerbate these difficulties given the low statistical power of most human neuroscience studies. I discuss potential approaches to addressing the challenges confronting clinical science with respect to replicability, including the use of analogue designs in nonclinical samples, the development of cross-laboratory collaborations that share a common protocol, and the need to "think meta-analytically" – to appreciate that each study is merely one sample of a population of potential studies that have yet to be conducted.
Electronic archiving:	Yes
Membership status:	Non-Member

Participant 6:	Ruthellen Josselson, PhD *Fielding Graduate University, Santa Barbara, CA*

E-Mail address:	rjosselson@fielding.edu
Mailing address:	4210 Tuscany Ct., Baltimore, MD 21210
Title of presentation:	Conversation as Replication in Qualitative Research
Summary:	Replication has quite different meanings in qualitative research, which is epistemologically based in an interpretive hermeneutics. Qualitative researchers assume that a study cannot be repeated in exactly the same way by someone else or even by the same person with different participants. All knowledge is situated. Any given research project reflects the position of the researcher in his or her culture or subculture and is influenced by gender, race, class, nationality, and other social identities including academic mooring in a discipline. Differences in findings about similar research questions become a source of intrigue rather than despair. Our effort is to systematically consider the relationships among research efforts to investigate interconnected phenomena of human life. Without null hypotheses to reject or effect sizes to compute, we instead are occupied with multifaceted historical and social contextualization; instead of direct relationships between motive and action, we concern ourselves with processes that are complex and multivalent. Qualitative research builds a thematic nomological net that relates findings and interpretations to one another, taking account of the circumstances in which the data were obtained and the person or people who are doing the study and the analysis. We take as a premise that good science is a narrative. That is, good science tells a story of how things are and tries to makes sense of phenomena. The challenge in qualitative research is aggregation, cumulative story development, meaningfully relating studies to one another. Conversation is the context in which knowledge is to be understood. The frontier of qualitative inquiry these days is the problem of creating and managing a conversation among us – that is, synthesizing and accumulating knowledge and understandings. This paper will suggest that the conversational approach to aggregation might be a better model for knowledge generation than rigid ideas about replication.
Electronic archiving:	Yes
Membership status:	APA Fellow

Participant 7:	Steven J Heine, PhD *University of British Columbia, Vancouver, BC, Canada*
E-Mail address:	heine@psych.ubc.ca
Mailing address:	Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver, BC V6T 1Z4 Canada
Title of presentation:	Replicability and WEIRDness: Thoughts on a Future Psychology
Summary:	While the replicability crisis represents a challenge for psychology of internal validity, a concomitant challenge for psychology is one of external validity; do our findings generalize to other contexts? Indeed, a relatively small proportion of psychological phenomena are investigated in multiple cultures, and one review found that psychology has the smallest proportion of citations coming from outside of the US out of 20 different scientific disciplines. This wouldn't be such a large problem if psychological phenomena appeared similarly around the world, however, many of them do not. Indeed, the findings for many key psychological phenomena from American undergraduates are often outliers in the context of the world's cultures. This external validity problem is relevant to the replicability crisis in two ways: First, failed replications that are conducted in different

	cultures may reflect the role of different cultural processes rather than a problem of internal validity; they may indicate boundary conditions. Second, this external validity problem points to a rarely discussed cost of one of the solutions to the replicability crisis: larger sample sizes. As samples increase in size so does the relative attractiveness of using convenience samples for conducting the research, such as American MTurk workers or American undergraduates. Larger sample sizes create disincentives for researchers to explore their phenomena with other less convenient samples. Implications and recommendations will be discussed.
Electronic archiving:	Yes
Membership status:	Non-Member

Participant 8:	Louise KW Sundararajan, PhD, EdD *Rochester Psychiatric Center, Rochester, NY*
E-Mail address:	louiselu@frontiernet.net
Mailing address:	691 French Road, Rochester, NY 14618
Phone numbers:	585-461-0995 (home)
Title of presentation:	From Replication to Translation: Fairness as Value for Cross-Cultural Research
Summary:	Replication across cultures is problematic even when it is successful. First, statistical significance may mask the lack of ecological validity. For instance, the stimulus used in the original study may not be representative of stimuli in a different culture; or the results may actually reflect affordances discovered by the local participants for the experimental tasks, rather than the putative psychological mechanisms they use in real life. Such issues concerning ecological validity are rarely raised when the p value reaches significance. Second, preoccupation with methods to enhance replication success may contribute to results that have more to do with the stimuli used for the testing, rather than the population sample being tested. The solution, from the perspective of indigenous psychology, lies in a paradigm shift from replication to translation. In the translation framework, contexts and differences are foundations for, instead of obstacles to, generalization. Instead of replicability, the generalization question can be handled by a translatability test, for which any local category can be used as a criterion. A scientific category needs to pass the translatability test of as many local categories as possible to qualify as a universal. In contrast to the conventional top down standardization process in the service of generalization, the translatability test is a bottom up process toward consensus building, the basis of which rests not on significance testing so much as on fair evaluation (Fiedler & Wäke, 2013).
Electronic archiving:	Yes
Membership status:	APA Fellow

Participant 9:	Brian Schiff, PhD *American University of Paris, Paris, France*
E-Mail address:	bschiff@aup.edu
Mailing address:	Department of Psychology, American University of Paris, 5 Boulevard de la Tour-Maubourg, Paris, France

Title of presentation:	Psychology's Silent Crisis
Summary:	The Open Science Collaboration's (OSC) findings, published in Science, that more than half of the research studies published in the most competitive and highly vetted journals did not demonstrate consistent results when the same design and procedures were repeated should be a catalyst for meaningful discussions on the foundations of psychological research. Proposals to encourage replication and make the process of research more visible and collaborative are innovative and should be embraced. But, psychology's methods problems are much more profound, systematic and enduring than the current controversy imagines. There is nothing inherently problematic about reductive statistical analysis--such methods are very useful tools for examining a wide range of research questions. But, variable centered research is a poor method for understanding persons or for describing how psychological processes function. Psychologists need to ask some tough questions about the entire research enterprise: What are the limits of variable centered methods for understanding human psychology? Even if they are repeatable, what is the value of our discoveries? Using the example of research on personality consistency or change over time, I describe the ubiquitous error that psychologists make interpreting statistical results derived from populations as evidence for individual psychological processes, what James Lamiell (2003) has called the Thorndike maneuver. In order to get beyond the crisis, psychological research needs to open up to innovative theories and methods for exploring human lives one by one. In this light, I argue for the utility of a narrative perspective and how narrative can make a contribution to the future of scientific psychology by directing us to the ongoing process of how persons interpret and re-interpret experience, self, other and the world.
Electronic archiving:	Yes
Membership status:	APA Member

Participant 10:	Lisa M Osbeck, PhD *University of West Georgia, Carrollton, GA*
E-Mail address:	losbeck@westga.edu
Mailing address:	Department of Psychology, University of West Georgia, 1601 Maple St., Carrollton, GA 30118
Title of presentation:	Failure to Replicate: Crisis or Chrysalis for Psychological Science?
Summary:	We analyze the meaning of psychology's replication crisis in the context of the broader project of scientific advancement. The metaphor of the chrysalis suggests something that may appear initially as a barrier or constraint, yet from a broader perspective is essential to development. Situating the "crisis" in the history of philosophy of science, we interpret the "new instrument" of scientific method as at base a set of practices for generalizing from experience in the face of variation and human fallibility. The principal tool is a procedure of meticulous comparison, with the corresponding rule to generalize cautiously and give equal attention to counter-instances. We argue that such comparative analysis applies to all empirical research, though the sources of variation and fallibility are extremely complex in human science. We suggest that failure to replicate experimental results is analogous to a counter-instance variation, one to be taken up and analyzed through comparison within the overall inductive project of the science. Therefore replication problems in psychology must be understood within the context of the role of "failure" in science more

	generally. Through comparative analysis of our own, we demonstrate how the fallibility inherent in empirical research renders it fundamentally different from the formal sciences of mathematics, geometry, and logic, in which certainty and perfect replication are achieved. Moreover, empirical replication failure functions not only to check or constrain generalization but enables discovery and facilitates new insights, which we illustrate with a case example from an ethnographic study of laboratory research.
Electronic archiving:	Yes
Membership status:	APA Fellow
Coauthor(s):	Frederick J Wertz, PhD, *Fordham University, Bronx, NY*

Edit **Cancel** **Print** **Save & Exit** **Submit**