Addressing Challenges to Replicability in Behavioral and Psychological Science
An APS, NIA & OBSSR Joint Meeting
June 4, 2012

Participants

Stefano Bertuzzi, National Institute of Mental Health
Daniel Kahneman, Princeton University
Robert Kaplan, NIH Office of Behavioral and Social Sciences Research
Jonathan King, National Institute on Aging
Alan Kraut, Association for Psychological Science
Lisbeth Nielsen, National Institute of Aging
Brian Nosek, University of Virginia
Hal Pashler, University of California, San Diego
David Reiss, National Institute on Aging /Yale University
Henry L. "Roddy" Roediger, III, Washington University in St. Louis
Jonathan Schooler, University of California, Santa Barbara
Barbara Spellman, University of Virginia
Richard Suzman, National Institute on Aging
Deborah Zarin, National Library of Medicine/ClinicalTrials.Gov

---

Background

Issues of research integrity and the replication of research findings are central concerns in all disciplines. These issues in the social and behavioral sciences have been receiving a great deal of recent attention. Current discussions have focused on strengthening the role of replication in advancing scientific knowledge, and recognizing and correcting false-positive results. Specific topics include:

- False-positive findings in research involving large scale data sets with multiple outcome variables
- Quality and appropriateness of data analysis techniques for large data sets
- Effect sizes for psychological phenomena
- A bias toward publishing positive results and the non-publication of negative results, leading to the "file-drawer" problem
- Access to data for purposes of replication by others

On June 4, 2012, a group of senior behavioral scientists and federal agency officials gathered to

1. Explore issues related to replication and reporting of false-positive results

2. Begin to formulate strategies for researchers, journal editors, and funding organizations that might remedy these problems

Following is a summary of the challenges and ideas that were identified by the group.

Challenges to Replication

The Nature of the Scientific Process

- There are legitimate reasons why researchers may not publish negative results or may change a study design or procedure midway through data collection.

- Changing a study procedure or stopping data collection mid-study when something is clearly not working reflects the dynamic process of research.

- Imposing too many restrictions on researchers may hamper the research process.

- Observational studies should be integrated with and utilized alongside experimental studies, interventional studies, and clinical trials to elucidate the pathways to health and disease.

Replication Success and Failure: Recognizing That Both Are Important

Replication studies can take one of two forms: the rarer *direct replication* (in which researchers attempt to replicate findings by matching the original experiment as closely as possible) and the more common *conceptual replication* (in which researchers try to recreate an experimental outcome using slightly different methodology).

- There are many reasons experimental findings may not replicate (e.g., "subtle phenomena" and a "decline effect—the tendency for effect sizes to diminish over time").

- Valid subtle phenomena, in particular, may be highly sensitive to seemingly minor variations in procedure and may take a level of skill to replicate that is not apparent in the study methods. Thus, subtle phenomena may be undesirable, but that does not necessarily make them illegitimate.

- Replication failures can help define the boundaries of psychological phenomena, but the current culture does not value replication research.

Role of researchers: Two major issues that may be contributing to problems of replication and poor data/analysis quality are lax procedures and inadequate management of controversy.

- Researchers are not always precise in creating, reporting and (by replicators) in exactly following experimental procedures.

- Failure to closely follow procedures makes successful replication difficult.

- Resistance on the part of original researchers (i.e., reluctance to share information or engage in discussions about points of controversy) impedes the collaboration necessary to facilitate replication.

Role of institutions: Current research funding and publication policies discourage replication studies.

- Replication studies do not bring as much prestige as "new" findings.
- Those wishing to perform replication research may have difficulty obtaining funding; current granting mechanisms do not include replication as a priority.
- Once replication studies have been completed, authors often have trouble publishing them because they are not novel or because their results are negative.

We believe that any serious science must value and encourage replication (and non-replication). Following are strategies to enhance replicability and replication efforts in behavioral and psychological science.

1. Research Repositories

    a. Require that researchers declare their variables, methods, and hypotheses prior to beginning their study in a registry along the model of ClinicalTrials.gov, for example. Taking this a step further, require that researchers specify this information as well as log their lab meeting notes, analysis procedures, datasets, etc. in an archiving system such as the Open Science Framework (http://openscienceframework.org/).

    b. To encourage participation, grant makers and journal editors could encourage (or even require) such registration and logging as a condition of funding or publication.

2. "Adversarial" Collaborations

    a. Establish a culture (through funding and publication policies) in which interaction between authors and serious critics is encouraged and rewarded.

    b. When designing a replication study, researchers should be encouraged to collaborate with those who originally conducted the research.

3. Replication Research

    a. Provide funding explicitly for replication research. This could be accomplished through the creation of replication centers, earmarking funds specifically for replication research, issuing contracts for replication studies, and/or other mechanisms.

    b. Consider funding a few replication centers. The mission of these centers would be to replicate selected studies and to report the results.

    c. Provide funding to support "adversarial" research collaborations.

d. Build an expectation for researchers--via funding agencies and their institutions--that they need to increase their involvement in replication efforts both by conducting replication studies and by cooperating with others' attempts to replicate their work.

e. Create awards to be used as incentives by funding agencies to encourage replication research.

f. Identify and prioritize areas where replication is most important.

4. Publication Policies

a. New journals or special issues/sections of existing journals dedicated to replication research should be considered.

b. Editors could create distinct journal sections for confirmatory and exploratory research and could give a special designation to studies that have been pre-registered or archived into a database such as the Open Science Framework.

c. Encourage a shift from the practice of publishing critiques and rejoinders and instead encourage collaboration when there are differing interpretations of data.

d. Undertake a small-scale internal study to explore the effect of journal reviewer checklists on increasing the quality of article review processes.

5. Data Collection and Analysis

a. Data-sharing policies such as NIH's should be developed and enforced by funding agencies, institutions, and journals to encourage data sharing and transparency among researchers. Researchers who use others' datasets should acknowledge the datasets' creators (perhaps via authorship).

b. Exploratory findings should be verified though confirmatory analysis before publication.

c. Researchers should explicitly note in their publications when they move from exploratory to confirmatory analysis.

d. Raising the accepted alpha (p) level could also help prevent publication of false-positive findings.

e. Strengthen and enforce data collection standards to facilitate replication.

Next Steps

The group agreed that the full day discussion was a valuable starting point for a conversation that might extend beyond and within and social and behavioral science and that additional work is required. The creation of small working groups, possibly organized along the categories listed above, was suggested to expand on the discussion and propose directions forward.

One more specific follow-up involves the journal *Psychological Science*, discussed as a high visibility journal where certain of these ideas might be tested. (*Psychological Science* editor Eric Eich was to attend the June 4 workshop, but was unable to at the last minute.) Following the meeting, Roediger and Kraut briefed Eich and Eich agreed to experiment in *Psychological Science* with a "logging of laboratory procedures" option for author submissions (i.e., a process by which experimenters would provide detailed documentation). A subgroup of Eich, Roediger, Pashler and Kraut are now engaged in discussing how to develop this so that it encourages more transparent reporting of procedures and data for others to access. The larger group will explore these policies with other journals.