



## The Replication Recipe: What makes for a convincing replication? <sup>☆</sup>



Mark J. Brandt <sup>a,\*</sup>, Hans IJzerman <sup>a,1</sup>, Ap Dijksterhuis <sup>b,2</sup>, Frank J. Farach <sup>c,2</sup>, Jason Geller <sup>d,2</sup>, Roger Giner-Sorolla <sup>e,2</sup>, James A. Grange <sup>f,2</sup>, Marco Perugini <sup>g,2</sup>, Jeffrey R. Spies <sup>h,2</sup>, Anna van 't Veer <sup>a,i,2</sup>

<sup>a</sup> Tilburg University, Netherlands

<sup>b</sup> Radboud University Nijmegen, Netherlands

<sup>c</sup> University of Washington, USA

<sup>d</sup> Iowa State University, USA

<sup>e</sup> University of Kent, UK

<sup>f</sup> Keele University, UK

<sup>g</sup> University of Milano-Bicocca, Italy

<sup>h</sup> Center for Open Science, USA

<sup>i</sup> TIBER (Tilburg Institute of Behavioral Economics), Netherlands

### HIGHLIGHTS

- Close replications are an important part of cumulative science.
- Yet, little agreement exists about what makes a replication convincing.
- We develop a Replication Recipe to facilitate close replication attempts.
- This includes the faithful recreation of a study with high statistical power.
- We discuss evaluating replication results and limitations of replications.

### ARTICLE INFO

#### Article history:

Received 10 July 2013

Revised 12 October 2013

Available online 23 October 2013

#### Keywords:

Replication

Statistical power

Research method

Pre-registration

Solid Science

### ABSTRACT

Psychological scientists have recently started to reconsider the importance of close replications in building a cumulative knowledge base; however, there is no consensus about what constitutes a convincing close replication study. To facilitate convincing close replication attempts we have developed a Replication Recipe, outlining standard criteria for a convincing close replication. Our Replication Recipe can be used by researchers, teachers, and students to conduct meaningful replication studies and integrate replications into their scholarly habits.

© 2013 The Authors. Published by Elsevier Inc. All rights reserved.

### Introduction

Replicability in research is an important component of cumulative science (Asendorpf et al., 2013; Jasny, Chin, Chong, & Vignieri, 2011; Nosek, Spies, & Motyl, 2012; Rosenthal, 1990; Schmidt, 2009), yet relatively few close replication attempts are reported in psychology (Makel, Plucker, & Hegarty, 2012). Only recently have researchers systematically reported replications online (e.g., [psychfiledrawer.org](http://psychfiledrawer.org),

[openscienceframework.org](http://openscienceframework.org)) and experimented with special issues to incorporate replications into academic publications (e.g., Nosek & Lakens, 2013; Zwaan & Zeelenberg, 2013). Moreover, some prestigious psychology journals (e.g., *Journal of Experimental Social Psychology*, *Journal of Personality and Social Psychology*, *Psychological Science*) are recently willing to publish both failed and successful replication attempts (e.g., Brandt, 2013; Chabris et al., 2012; LeBel & Campbell, in press; Matthews, 2012; Pashler, Rohrer, & Harris, in press) and even devote ongoing sections to replications (see the new section in *Perspectives on Psychological Science*, Registered replication reports, 2013).

From initial conclusions drawn from replication attempts of important findings in the empirical literature, it is clear that replication studies can be quite controversial. For example, the failure of recent attempts to replicate “social priming” effects (e.g., Doyen, Klein, Pichon, & Cleeremans, 2012; Pashler et al., in press) has prompted

<sup>☆</sup> This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

\* Corresponding author.

E-mail address: [m.j.brandt@tilburguniversity.edu](mailto:m.j.brandt@tilburguniversity.edu) (M.J. Brandt).

<sup>1</sup> First two authors share first authorship.

<sup>2</sup> All other authors share second authorship.

psychologists and science journalists to raise questions about the entire phenomenon (e.g., Bartlett, 2013). Failed replications have sometimes been interpreted as 1) casting doubt on the veracity of an entire subfield (e.g., candidate gene studies for general intelligence, Chabris et al., 2012); 2) suggesting that an important component of a popular theory is potentially incorrect (e.g., the status-legitimacy hypothesis of System Justification Theory, Brandt, 2013); or 3) suggesting that a new finding is less robust than when first introduced (e.g., incidental values affecting judgments of time; Matthews, 2012). Of course, there are other valid reasons for replication failures: Chance, misinterpretation of methods, and so forth.

Nevertheless, not all replication attempts reported so far have been unsuccessful. Burger (2009) successfully replicated Milgram's famous obedience experiments (e.g., Milgram, 1963), suggesting that when well-conducted replications *are* successful they can provide us with greater confidence about the veracity of the predicted effect. Moreover, replication attempts help estimate the effect size of a particular effect and can serve as a starting point for replication–extension studies that further illuminate the psychological processes that underlie an effect and that can help to identify its boundary conditions (e.g., Lakens, 2012; Proctor & Chen, 2012). Replications are therefore essential for theoretical development through confirmation and disconfirmation of results. Yet there seems to be little agreement as to what constitutes an appropriate or convincing replication, what we should infer from replication “failures” or “successes,” and what close replications mean for psychological theories (see e.g., the commentary by Dijksterhuis, 2013 and the reply by Shanks & Newell, 2013). In this paper, we provide our Replication Recipe for conducting and evaluating close replication attempts.

### Close replication attempts

In general, how can one define close replication attempts? The most concrete goals are to test the assumed underlying theoretical process, assess the average effect size of an effect, and test the robustness of an effect outside of the lab of the original researchers by recreating the methods of a study as faithfully as possible. This information helps psychology build a cumulative knowledge base. This not only aids the construction of new, but also the refinement of old, psychological theories. In the definition of our Replication Recipe, close replications refer to those replications that are based on methods and procedures *as close as possible* to the original study. We use the term *close replications* because it highlights that no replications in psychology can be absolutely “direct” or “exact” recreations of the original study (for the basis of this claim see Rosenthal, 1991; Tsang & Kwan, 1999). By definition then, close replication studies aim to recreate a study as closely as possible, so that ideally the only differences between the two are the inevitable ones (e.g., different participants; for more on the benefits of close replications see e.g., Schmidt, 2009; Tsang & Kwan, 1999).

### The Replication Recipe

What constitutes a convincing close replication attempt, and how does one evaluate such an attempt? This is what the Replication Recipe seeks to address. The Replication Recipe is informed by the goals of a close replication attempt: Accurately replicating methods and estimating effect sizes and evaluating the robustness of the effect outside the lab of origin. Our discussion is based on a synthesis of our own trials and errors in conducting replications and guidelines recently developed for special issues and sections of psychology journals (Nosek & Lakens, 2013; Open Science Collaboration, 2012; Registered replication reports, 2013; Zwaan & Zeelenberg, 2013). In this synthesis, we make explicit the expectations and necessary qualities of a convincing replication that can be used by researchers, teachers, and students when designing and carrying out replication studies.

A convincing close replication *par excellence* is executed rigorously by independent researchers or labs and includes the following five additional ingredients:

1. Carefully defining the effects and methods that the researcher intends to replicate;
2. Following as exactly as possible the methods of the original study (including participant recruitment, instructions, stimuli, measures, procedures, and analyses);
3. Having high statistical power;
4. Making complete details about the replication available, so that interested experts can fully evaluate the replication attempt (or attempt another replication themselves);
5. Evaluating replication results, and comparing them critically to the results of the original study.

Each of these criteria is described and justified below. We present and explain 36 questions that need to be addressed in a solid replication (see Table 1<sup>3</sup>). This list of questions can be used as a checklist to guide the planning and communication of a study and will help readers and reviewers to evaluate the replication, by understanding the decisions that a replicator has made when designing, conducting, and reporting their replication. These questions are intended to help replicators follow the Replication Recipe and determine when and why they have deviated from the five Replication Recipe ingredients.

*Ingredient #1: Carefully defining the effects and methods that the researcher intends to replicate*

Prior to conducting a replication study, researchers need to carefully consider the precise effect they intend to replicate (Questions 1–9), including the size of the original effect (Question 3), the effect size's confidence intervals (Question 4) and the methods used to uncover it (Questions 5–9). Although this can be a straightforward task, in many studies the effect of interest may be a specific aspect of a more complicated set of results. For example, in a 2 × 2 design where the original effect was a complete cross-over interaction, such that an effect was positive in one condition and negative in the other, the effect of interest may be the interaction, the positive and negative simple effects, or perhaps just one of the simple effects. On other occasions, the information about the methods used to obtain the effect will be unclear (e.g., the original country the study was completed in, Question 7); in these cases, it may be necessary to ask the original authors to provide the missing information or to make an informed guess. It is important to know the precise effect of interest from the beginning of the design-phase of the replication because it determines nearly all of the decisions that follow. A related consideration, especially when resources are limited, is the importance and necessity of replicating a particular effect (Question 2). Such decisions to replicate or not should be based on either the effect's theoretical importance to a particular field or its direct or indirect value to society. Another consideration is existing confidence in the reliability of the effect; an effect with a number of existing close replications in the literature may be less urgent to replicate than one without any such support (see discussion of the [Replication value project, 2012–2013](#)). In other words, not every study is worth replicating. By considering the theoretical and practical importance of a finding the best allocation of resources can be made.

*Ingredient #2: Following exactly the methods of the original study*

Once a study has been chosen for replication, and the precise effect of interest has been identified, the design of the replication study can commence. In an ideal world, the methods of the original study

<sup>3</sup> Also available as a pre-registration form on [open science.org](http://open science.org)

Download English Version:

<https://daneshyari.com/en/article/7324956>

Download Persian Version:

<https://daneshyari.com/article/7324956>

[Daneshyari.com](https://daneshyari.com)