Journal of Experimental Social Psychology xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

On the scientific superiority of conceptual replications for scientific progress

Christian S. Crandall [a,*], Jeffrey W. Sherman [b]

[a] University of Kansas, USA
[b] University of California, Davis, USA

ARTICLE INFO

Article history:
Received 1 November 2014
Revised 1 October 2015
Accepted 11 October 2015
Available online xxxx

Keywords:
Replication
Conceptual replications
Social psychology
Philosophy of science

ABSTRACT

There is considerable current debate about the need for replication in the science of social psychology. Most of the current discussion and approbation is centered on *direct* or exact replications, the attempt to conduct a study in a manner as close to the original as possible. We focus on the value of *conceptual* replications, the attempt to test the same theoretical process as an existing study, but that uses methods that vary in some way from the previous study. The tension between the two kinds of replication is a tension of values—exact replications value confidence in operationalizations; their requirement tends to favor the status quo. Conceptual replications value confidence in theory; their use tends to favor rapid progress over ferreting out error. We describe the many ways in which conceptual replications can be superior to direct replications. We further argue that the social system of science is quite robust to these threats and is self-correcting.

> Any working scientist is more impressed with 2 replications in each of 6 highly dissimilar experimental contexts than he is with 12 replications of the same experiment. (Meehl, 1990, p. 111.)

Scientific ideas must be robust before they can be endorsed. They must be testable and they must inspire the confidence of a skeptical audience. There are many ways that these ideas acquire a reliable place in the marketplace of ideas: elegance, intuitiveness, explanatory power, rigor, and so on. But the ability of a phenomenon to be replicated is a necessary condition for widespread acceptance by scientists (Schmidt, 2009). Catchy, interesting, telling, and surprising results can all have currency, but if an effect proves unreliable or impossible to replicate, support for the idea will not—must not—persist.

> Only when certain events recur in accordance with rules or regularities, as is the case with repeatable experiments, can our observations be tested—in principle—by anyone. We do not take even our own observations quite seriously, or accept them as scientific observations, until we have repeated and tested them. (Popper, 1959, p. 23).

Time and chance occur for every data set—it is only through replication that we can be confident. Over the past few years, a concern about the repeatability and replicability of experiments has spread throughout the social psychological community. In its wake, a variety of articles, blog posts, research programs, and non-profit organizations have come forward with prescriptions for a more replicable science of social psychology (e.g., Open Science Collaboration, 2012). Some of this work has been well funded and well covered by science journalists (e.g., Meyer & Chabris, 2014). There is a broad consensus in favor of robust findings, for reliability in the scientific record, for high quality research with dependable reporting and replicability, and for progress in scientific knowledge. But there are sharp differences among scientists in (1) which scientific goals should take priority over others and (2) the best way to meet those respective goals.

1. Two kinds of replications

One of the most prominent fault lines among scientists is in their approaches to replication. Following standard discourse, we divide replication into two kinds—exact replication and conceptual replication. An *exact* or *direct* replication is an attempt to conduct a study, usually published in a peer-reviewed journal, in a manner as close to the original as possible. An exact replicator seeks to use the same materials, the same manipulations, the same dependent variables, and the same kind of participants as the originally published article. A replication is considered "successful" when the exact replication results in a pattern of data that mimics—or is close to—the original article's findings.

The second class of replications is known as *conceptual* replications. A conceptual replication is an attempt to test the same fundamental idea or hypothesis behind the original study, but the operationalizations of the phenomenon, the independent and dependent variables, the type and design of the study, and the participant population may all differ substantially. (Others have called this a distinction between replicability [exact replication] and repeatability [conceptual

* Corresponding author.

http://dx.doi.org/10.1016/j.jesp.2015.10.002
0022-1031/

replication], e.g., Casadevall & Fang, 2010; Drummond, 2009) The critical difference between an exact and a conceptual replication is whether or not they share the same operationalizations of the theory. Although this distinction is common in discussions among social psychologists, the scientific and philosophical literature on the matter is surprisingly scant (cf. McGrath, 1981).

In some fields of science, an exact replication is a sensible proposition. In physics, chemistry, biology or geology, the processes that affect an outcome are usually transhistorical and transcultural—language, politics, and social history rarely affect gravity, electron weight, the structure of proteins, or water flow through limestone. The meaning of the operationalization is consensual among scientists; the value of mass or acceleration, bone density, or the Mohs hardness scale, for example.

But in matters of social psychology, one can never step in the same river twice—our phenomena rely on culture, language, socially primed knowledge and ideas, political events, the meaning of questions and phrases, and an ever-shifting experience of participant populations (Ramscar, 2015). At a certain level, then, all replications are "conceptual" (Stroebe & Strack, 2014), and the distinction between direct and conceptual replication is continuous rather than categorical (McGrath, 1981). Indeed, many direct replications turn out, in fact, to be conceptual replications. At the same time, it is clear that direct replications are based on an attempt to be as exact as possible, whereas conceptual replications are not.[1]

2. Replications and theoretical consequences

> *The meaning of theoretical terms cannot be totally exhausted by operational definitions, but the ways in which theoretical terms function in science cannot be understood in the absence of the ways in which they are operationalized.* (Hull, 1988, pp. 516–517).

There is no controversy over the need for replication; virtually all scientists and philosophers of science endorse the notion that replication of one sort or another is absolutely essential. The controversy is largely over the degree to which different kinds of replications advance scientific knowledge. Historically, research in psychology has favored conceptual over direct replications. Most researchers were trained to value the pursuit of robustness and generality of theoretical ideas over the repeatability of a particular study. However, recent observations that direct replications may reproduce original findings at a lower rate than expected have led to calls for increasing the frequency and publication of direct replications (Open Science Collaboration, 2015). At the same time, conceptual replications have been increasingly criticized for biasing research toward confirmation and impeding the possible disconfirmation of research findings and the theories they support (e.g., Nosek, Spies, & Motyl, 2012; Pashler & Harris, 2012; Roberts, 2014).

In terms of both published commentary (Pashler & Wagenmakers, 2012) and public social media discussions on the "replication crisis," those favoring a shift toward direct replication have received a great deal of attention. As described above, it is certainly true that, historically, the potential advantages of direct replication have received relatively little attention. As such, the recent discussions can be seen as an important corrective to a historical imbalance. At the same time, the virtues of conceptual replication have been overlooked or directly challenged in these venues. Given the field's historical predilection toward conceptual replication, this may be of little consequence. However, it is our experience that psychological researchers, particularly those in the early stages of training, are increasingly prone to dismiss the potential benefits of conceptual replication in favor of direct replication. Moreover, the recent focus on direct replication seems to have created the perception of consensus that direct replications are of greater value than conceptual replications. Again, we have noticed these tendencies particularly among younger members of our research guild. In this context, we believe that it is important to provide a corrective and to articulate the importance of conceptual replication. The primary purpose of this paper is to do that and, at the same time, offer suggestions to increase the scientific value of conceptual replication.

3. The purposes of direct and conceptual replication

Whereas direct replications enhance one's confidence in operationalizations, conceptual replications enhance one's confidence in theoretical hypotheses. In 1906, the physicist Pierre Duhem (1906/1954) pointed out that every empirical scientific test was comprised of a conjunction of the theoretical hypothesis and its operationalizations—every empirical test conflates *ideas* with *methods*. A "failure" of an empirical test is always ambiguous, because the failure may indicate that the idea is incorrect (e.g., a failed conjecture or "falsification," Popper, 1963/2002), or it may indicate that the operationalization process failed, or both. Exact replications can never speak to this ambiguity, they can only perpetuate it; this makes straightforward falsification a logical impossibility (see also Meehl, 1990; Quine, 1980).

Conceptual replications disperse this ambiguity, and as a result, can contribute more to theoretical development and scientific advance. If an idea replicates *across* operationalizations, then the idea is substantially more likely to be correct than if it replicates using the exact same operationalizations, no matter how many times or with whatever precision. As such, conceptual replications are critical for establishing the *generalizability* of an initial observation and the theory it purports to support.

The history of science is replete with examples in which an original demonstration is met with substantial skepticism, with specific complaints about confounds, alternative explanations, or concerns about the effectiveness of methods. In these cases, direct replications meet with exactly the same complaints, but conceptual replications can prove far more persuasive. Mackay and Oldford (2000) reviewed critical studies about whether the speed of light was infinite or merely very, very fast. In 1671, Ole Rømer measured anomalies in the timing of eclipses of Io, one of Jupiter's moons, and showed that these anomalies could be accounted for by the distance between Earth and Jupiter; when Earth and Jupiter were closer, the eclipse came sooner. Because Rømer knew the size of the Earth's orbit, he was able to make a fairly accurate estimate of the speed of light. But this research did not convince many of the leading scientists of the day, including Descartes and Giovanni Cassini, the director of the Paris Observatory, who offered many criticisms and alternative explanations for the data, despite the fact that Rømer replicated his findings over many years (Soter & deGrasse Tyson, 2000). It wasn't until 1729, when James Bradley used the parallax motion of stars (because stars are different distances from Earth, the Earth's movement causes an apparent shifting of the stars' relative position) to calculate the speed of light, and calculated a very similar number to Rømer's. Following the use of this technique the scientific community agreed that the speed of light was not infinite, but rather about 300,000,000 m per second.

Exact replications are often unconvincing to the scientific community—skeptics require a different method for test of the same hypothesis. The second operationalization often dispels spurious criticisms of the first study's method (MacKay & Oldford, 2000). Exact replication could create precision in estimation, but it would not convince the scientific community of *the meaning of the observation*; a conceptual replication did.

[1] If one prefers the continuous approach, the reader might interpret this paper to favor movement in the direction of greater conceptual replication and away from more exact replications.

Please cite this article as: Crandall, C.S., & Sherman, J.W., On the scientific superiority of conceptual replications for scientific progress, *Journal of Experimental Social Psychology* (2015), http://dx.doi.org/10.1016/j.jesp.2015.10.002