



# Psychological Inquiry

An International Journal for the Advancement of Psychological Theory

ISSN: (Print) (Online) Journal homepage: <https://www.tandfonline.com/loi/hpli20>

## There's Nothing Social about Social Priming: Derailing the "Train Wreck"

Jeffrey W. Sherman & Andrew M. Rivers

To cite this article: Jeffrey W. Sherman & Andrew M. Rivers (2021) There's Nothing Social about Social Priming: Derailing the "Train Wreck", *Psychological Inquiry*, 32:1, 1-11, DOI: [10.1080/1047840X.2021.1889312](https://doi.org/10.1080/1047840X.2021.1889312)

To link to this article: <https://doi.org/10.1080/1047840X.2021.1889312>



Published online: 01 Apr 2021.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)

## There's Nothing Social about Social Priming: Derailing the "Train Wreck"

Jeffrey W. Sherman<sup>a</sup> and Andrew M. Rivers<sup>b</sup>

<sup>a</sup>Department of Psychology, University of California, Davis, California, USA; <sup>b</sup>Department of Psychology, University of British Columbia, Vancouver, Canada

### ABSTRACT

Failures to replicate high-profile priming effects have raised questions about the reliability of so-called "social priming" phenomena. However, not only are many of the relevant studies not particularly social in nature, but other robust priming effects that are clearly social in nature do not count as social priming. Most importantly, the focus on the supposedly social aspect of the work has obscured factors that help to account for the relative reliability of priming effects. Here, we examine the construct of social priming, describe some simple demonstrations on the role of experimental design in priming reproducibility, and discuss future avenues for building a better understanding of priming. We conclude that the term "social priming" should be laid to rest, and that it is time to move past arguments about the reliability of specific effects and shift our energy to building theories that help us better understand the mechanisms underlying priming effects.

### KEYWORDS

Priming, social priming, train wreck

In this paper, we examine the controversy surrounding the replicability of "social priming" research.<sup>1</sup> We place the controversy in a historical context, review the construct of social priming, evaluate alternative characterizations of the relevant literature, describe some simple demonstrations that help clarify the nature of the controversy, and discuss future avenues of research for better understanding priming phenomena.

One could reasonably ask whether another examination of social priming is really necessary. Indeed, our own interest in publishing our analysis had waned. Social priming has been discussed widely, including in a special issue wholly devoted to it in the journal *Social Cognition* in 2014. Keith Payne and his colleagues made some of our central points in an excellent paper in 2016 (Payne, Brown-Iannuzzi, & Loersch, 2016). Moreover, we have been assured repeatedly that our main arguments in challenging common critiques of social priming are obvious and that everyone already understands them. Indeed, the issues are not complicated, and it is unlikely that critics of social priming will believe they have learned anything from reading this article.

Nevertheless, in 2017, at the same time we were being reassured, a number of analyses appeared in blogs and mainstream media that repeated the standard criticisms without acknowledging these issues (e.g., Engber, 2017; McCook, 2017; Schimmack, Heene, & Kesavan, 2017). Nor has the term disappeared from empirical academic research (e.g., Gilder & Heerey, 2018; Meyer, 2019). In December 2019, *Nature* published a news feature on social priming

(Chivers, 2019; Sherman & Rivers, 2020), with an associated podcast that, once again, promoted the same flawed narrative and commentary. That article included the greatest hits from the onset of the controversy, with a re-telling of Bem, Stapel, failed replications, and Kahneman's infamous letter. There was still no acknowledgement of simple definitional and methodological problems with the critique. This misrepresentation is not inconsequential. Beyond simply being wrong, it draws attention toward a handful of controversial effects and away from developing a deeper theoretical understanding of priming.

### The Looming Train Wreck

So, how did we get here? How did social priming become such a lightning rod of controversy? The year 2011 was eventful for scientific psychology, particularly social psychology. That year saw the publication in *Journal of Personality and Social Psychology* of Bem's article claiming to show scientific support for the operation of extrasensory perception (Bem, 2011). Later that same year, the social psychologist Diederik Stapel was found to have engaged in research fraud on a massive scale. Some of Stapel's fraudulent publications reported results from social priming studies. Also in 2011, Simmons, Nelson, and Simonsohn (2011) published their paper on the use of researcher degrees of freedom to produce statistically significant research results at will, highlighted by the absurd example of causing people to report a later birthdate after having been primed with the

Beatles song “When I’m 64.” This amusing demonstration bore superficial resemblance to social priming studies. The following year, Doyen, Klein, Pichon, and Cleeremans (2012) reported a failure to replicate a priming study published by social psychologists Bargh, Chen, and Burrows (1996), in which people primed with words associated with aging subsequently walked more slowly. The details surrounding this replication effort received considerable attention in an article by Ed Yong (Yong, 2012) that had been shared over 900 times on social media by August, 2017 (Discover magazine subsequently stopped publishing figures on social media sharing) and has been cited in academic literature more than 30 times. Subsequent failures to replicate other social priming effects also were widely discussed (e.g., Harris, Coburn, Rohrer, & Pashler, 2013; Shanks et al., 2013).

All of this contributed to the onset of an intense period of reexamination of research practices among experimental psychologists, commonly referred to as a “crisis of confidence” or a “supposed crisis of confidence,” depending on one’s perspective (Spellman, 2015). A particularly influential voice in instigating this movement belonged to Daniel Kahneman, the Nobel Prize winning cognitive psychologist. Kahneman had described a number of priming studies in his best-selling book *Thinking Fast and Slow* (again, 2011), and was embarrassed that researchers were having trouble replicating some of them. In 2012, he published an open letter to researchers working in what he referred to as the “field” of social priming, warning of a “train wreck looming,” and that the field had become the “poster child for doubts about the integrity of psychological research” (Kahneman, 2012). This letter attracted considerable professional and media attention, much of which consisted of critiques of social priming. A Google search of “Kahneman Train Wreck” yielded more than 350,000 results at the time of this writing, and the letter has been cited more than 100 times in academic journal articles.

Aside from accelerating introspection about research practices, Kahneman’s letter had two other significant impacts. First, it identified social psychological research as a particularly troubled field. Given his stature, it is not surprising that Kahneman’s prophecy about the field becoming a poster child was self-fulfilling. Second, in calling into question the integrity of the research, the letter contributed to the still-ongoing conflation of research fraud with reproducibility issues, particularly as it relates to social psychology (e.g., Neuroskeptic, 2017; Wagenmakers, 2014). This conflation was magnified by the Levelt Report (2012), which summarized the official investigation into Stapel’s fraudulent activities. The Report stated that “far more than was originally assumed, there are certain aspects of the discipline itself that should be deemed undesirable or even incorrect from the perspective of academic standards and scientific integrity.” This elicited much protest among social psychologists, including a rebuke from the European Association of Social Psychology.

All of these events help to explain the fixation on social priming, specifically, and social psychology, more generally,

as the Problem Child of psychological science—the appearance of Bem’s publication in a social psychological journal, Stapel’s status as a social psychologist, Simmons et al.’s parody of social priming, the commotion surrounding Doyen et al. (2012) failure to replicate Bargh et al. (1996), Kahneman’s letter (2017), and the Levelt Report (2012).

## The Invention of Social Priming

Broadly, priming refers to the phenomenon whereby exposure to a stimulus influences subsequent behavior without conscious guidance or intention. Frequently, priming effects are thought to result from the activation of mental representations that facilitate or interfere with related subsequent behavior (Molden, 2014). Research traditions have utilized priming tasks to investigate many different theoretical questions of interest. Cognitive psychologists initially used priming tasks to probe the organization of mental representations (e.g., Neely, 1991). Social psychologists initially adapted priming methodologies to understand how activated knowledge and evaluations influence perception and behavior (e.g., Fazio, Sanbonmatsu, Powell, & Kardes, 1986). At a more granular level, different areas of investigation have developed unique paradigms to elicit priming effects and have generated a plethora of theoretical models to explain those effects.

Across research traditions, priming involves exposing participants to stimuli—or “primes”—that are incidental to, and yet still influence, subsequent behavior. For example, work with the Lexical Decision Task (Meyer & Schvaneveldt, 1971) demonstrates that people are faster to correctly identify words (e.g., “doctor”) following the presentation of a related prime word (e.g., “nurse”) compared to an unrelated prime word (e.g., “paper”). In other words, the target behavior—correctly indicating that “doctor” is a word—occurs more quickly when related, incidental stimuli are observed than when unrelated stimuli are observed.

What, then, is social priming? The truth is that it did not exist as a Thing until the controversy that created it. In 2018, Neuroskeptic (Neuroskeptic, 2018) conducted a deep dive into the origins of the term “social priming.” They identified chapters in an edited volume (Bargh, 2005; Trope & Fishbach, 2005) as the first “modern” uses of the term. However, in both cases, the “social” in social priming was an adjective specifying the type of content that was being primed. “Social” modified the noun “priming,” but “social priming” was not its own Thing. In Neuroskeptic’s timeline, the first piece to name social priming as a Thing was a review article published in 2012 by Prabhakaran and Gray (2012), neither of whom are social psychologists (the scientists presumably conducting social priming research). This review described “the social priming paradigm,” but failed to include some of the earliest studies and most widely used priming measures that might have been labeled social priming. This was a troubling omen.

Though Neuroskeptic’s (2017) etymological effort was constructive and informative, it failed to identify what we (and we believe most social psychologists) consider to be the

true source of the term “social priming,” at least in terms of introducing it to a broad audience. The source for most of us was the open letter Kahneman wrote in the fall of 2012 that identified social priming as a “field” that was particularly problematic. At the time, there was considerable puzzlement among social psychologists, including those conducting the so-called social priming research, as to the meaning and origin of the term “social priming.” The senior author of the present paper had been Editor at the journal *Social Cognition* for 7 years at that time, and had never encountered the term. Likewise, many social psychologists working in the broad field of social cognition, who were quite familiar with the relevant studies, shrugged confusedly. Whatever the exact timeline of who said what and when, a few things are clear. First, at the time of Kahneman’s letter, social priming was not a recognized specific methodology or paradigm, never mind a field, as he had declared. Second, the idea of social priming as a Thing was not invented by the social psychologists largely doing and reading the work. Third, though the idea remains popular (the online title of Chivers, 2019 piece in *Nature* refers to the “field of social priming”), the term still has not been defined to anyone’s satisfaction.

### What Counts as Social Priming?

A major impediment to evaluating social priming is that there is no clear definition of what it is. Most obviously, it would seem that social priming should involve priming studies that include stimuli and/or responses related to social cognition and behavior. That is the spirit in which Bargh and Trope/Fishbach had used “social” as a modifier of “priming.” However, it was immediately apparent that the inclusion of social content was neither necessary nor sufficient for acquiring the label of social priming as envisioned by Kahneman and other critics of the research. Even the archetypal social priming paper (Bargh et al., 1996) fails to meet this criterion. There is nothing particularly social about the knowledge of a relationship between aging and walking speed or the notion that the salience of that knowledge could affect actual walking speed.

Perhaps more revealing than what is counted as social priming is what isn’t counted. There are huge bodies of research on the social psychology of attitudes, prejudice, and stereotyping that rely on priming paradigms. The Evaluative Priming Task (Fazio, Jackson, Dunton, & Williams, 1995) and the Weapons Identification Task (Payne, 2001) rely on standard sequential priming designs. Other frequently used indirect measures of attitudes and intergroup bias, such as the Implicit Association Test (Greenwald, McGhee, & Schwartz, 1998), the Go/No-Go Association Task (Nosek & Banaji, 2001), the First Person Shooter Task (Correll, Park, Judd, & Wittenbrink, 2002), the Affect Misattribution Procedure (Payne, Cheng, Govorun, & Stewart, 2005), and the Stereotype Misperception Task (Krieglmeyer & Sherman, 2012) rely on the same underlying logic of priming information in memory prior to measuring some associated behavioral response. All of these measures are clearly priming tasks, they were created by social psychologists, and they were

first reported in social psychology journals. In fact, there are orders of magnitude more published articles using these measures than the sorts described and critiqued by Kahneman and others. Moreover, these measures produce robust results that have been widely replicated (Cameron, Brown-Iannuzzi, & Payne, 2012). Upon being confronted with this large body of robust priming effects in the social psychological literature, many critics of social priming declared that, for a variety of reasons, these effects did not count as social priming. This makes clear that social priming is not simply priming research using social content or priming research conducted by social psychologists. Obviously, this contradiction introduces a major complication for a broad critique of a so-called paradigm or field of social priming.

### Behavioral Priming?

The most commonly stated basis for excluding these effects from social priming is that they do not examine the influence of primes on *behavior*. In this view, there is an important qualitative difference between the button presses in these paradigms and the behaviors reflected in social priming studies (e.g., Pashler, Rohrer, & Harris, 2013; Weingarten et al., 2016). To reflect this distinction, many researchers adopted the term “behavioral priming” to refer to the studies identified as social priming (e.g., Doyen et al., 2012; Wagenmakers, 2014). Though this tacit recognition of robust priming effects in the broader social psychological literature was welcome, the conversation surrounding the supposed problems with “social priming” continues unabated (e.g., Chivers, 2019; Engber, 2017; McCook, 2017; Schimmack et al., 2017). Often, the terms “social priming” and “behavioral priming” are used interchangeably, particularly outside of academic publishing, among authors who have accurately detected that the two terms refer to the same body of research (e.g., Chivers, 2019).

Rather than clarifying the contradictions inherent in the usage of the term “social priming,” the term “behavioral priming” introduces a new set of conceptual problems. What constitutes “behavior?” Which behaviors count? Pressing buttons on computer keyboards to judge the identity or valence of a target word or image is certainly behavior. What if the buttons being pushed are meant to represent shooting a gun (e.g., Correll et al., 2002)? What about using mouse clicks to choose gambles (Payne et al., 2016)? Judging the pleasantness of Chinese symbols (Payne et al., 2005) or the threat level of ambiguous faces (Krieglmeyer & Sherman, 2012)? Are these behaviors qualitatively different kinds of behaviors than circling a number on a rating scale (Srull & Wyer, 1979), filling in responses on a multiple-choice test of general knowledge (Dijksterhuis & van Knippenberg, 1998), rating the humor of cartoons (Strack, Martin, & Stepper, 1988) or completing word fragments (Zhong & Liljenquist, 2006), all of which are measures in studies commonly recognized as bone fide social priming? The point is that, just as no one ever provided a clear definition of social priming, no one has provided a coherent definition of behavioral priming. There is no theoretical

rationale for including some behaviors and excluding others. Words such as “broad,” “complex,” and “higher-order” are invoked, but they do not seem apt for the behavioral measures from social priming studies described above. We have yet to see any semblance of a cogent explanation of exactly when, why, and how a behavior crosses the threshold from being just a behavior to being a Behavior. As such, the behavior/non-behavior divide cannot capture the distinction between social psychological priming research that has been admitted and excluded from the category of social priming. We will return to this issue below.

### Long-Term versus Short-Term Priming

Wentura and Rothermund (2014) recognized an important distinction between priming paradigms that separate prime and target by milliseconds versus paradigms that separate them by multiple seconds or longer. This may serve as another basis for excluding robust priming paradigms in social psychology from counting as social priming (Wentura and Rothermund do not advocate this use of the distinction). Many paradigms that employ very brief gaps between prime and target have been demonstrated to be robust and replicable, whereas the controversial paradigms labeled as social priming generally impose much longer gaps between prime and target. Yet, the long-term/short-term distinction also does not capture the separation between research designated as social priming/not social priming. In particular, there are many examples of long-term priming that are not classified as social priming (e.g., Becker, Moscovitch, Behrmann, & Joordens, 1997; Hughes & Whittlesea, 2003; Sherman & Jordan, 2011; Tse & Neely, 2005, 2007; Woltz & Was, 2006, 2007).

### The Role of Experimental Design

To this point, we have determined that social priming is not defined by the presence of social content in a priming task, the home discipline of the researchers, or the length of time between prime and target. We also have argued that type of behavior does not provide a coherent basis upon which to distinguish studies that have been declared social priming/not social priming. What, then, best characterizes the distinction between studies that have been included or excluded as examples of social priming?

In our reading of the literature, the feature that seems to most closely covary with the social/nonsocial priming distinction is the status of a study as using a within- versus between-subjects design. Specifically, the studies identified as social priming, whether they include social content or not and regardless of the gap between prime and target, almost uniformly use between-subjects designs, whereas studies considered to exemplify nonsocial priming almost uniformly use within-subjects designs.<sup>2</sup>

<sup>2</sup>We use the term “almost exclusively” to acknowledge the fact that we are not familiar with every relevant published study. However, we are not aware of any studies consensually regarded as social priming that use a within-subjects design.

In between-subjects designs, each participant responds to a single type of prime, whereas in within-subjects designs, each participant responds to all prime-target combinations, usually over hundreds of trials. All else being equal, between-subjects designs offer substantially less statistical power than within-subject designs to detect the same effect. During the period in which much of the initial social priming research was conducted, little attention was paid to selecting sample sizes appropriate to study design. As such, the between-subjects social priming studies were frequently severely underpowered. For example, Bargh et al. (1996) elderly-priming studies used a fully between-subjects design with two conditions and one critical measurement per participant. Those studies sampled from 15 participants per condition. This would require an assumed effect size of  $d_s = 1.06$  to achieve Cohen’s (1988) recommended level of power,  $1 - \beta = 0.8$ . According to Cohen’s descriptions, this effect would be “grossly perceptible.” Given the subtlety of Bargh et al. (1996) manipulation, such an effect size seems unlikely.

Underpowered studies result in lower positive predictive value (PPV); that is, the probability that a “positive” research finding reflects a true effect (Button et al., 2013). Moreover, positive results along with low PPV are likely to provide exaggerated estimates of effect size. Thus, positive findings with low PPV are less likely to subsequently replicate than are positive findings with high PPV. In turn, selective publication policies favor the publication of positive research findings, while largely neglecting PPV. Thus, to the extent that published social priming studies were relatively more likely to be underpowered compared to other published priming studies, they also would be relatively less likely to subsequently replicate. It is not surprising that the most controversial and difficult to replicate priming effects have used between-subjects design, were inadequately-powered, and often had lower PPV (e.g., Bargh et al., 1996; Dijksterhuis & van Knippenberg, 1998). In contrast, the highly replicable priming effects in social psychology described above employ within-subjects designs in studies that are well-powered, and often had high PPV. Clearly, it is not the social versus nonsocial nature of the priming or the behavioral versus non-behavioral nature of the measures that is relevant in this comparison. Simply, adequately powered and published priming results are more likely to replicate than inadequately powered and published priming results.

Let us return to the origin of the term “social priming:” Kahneman’s embarrassment at having described priming studies that were difficult to replicate. In the fourth chapter of *Thinking Fast and Slow* (Kahneman, 2011), he describes the type of priming effects he later labeled as “social priming” and lamented as unreliable. Not all of the 29 social priming effects described in that chapter (see Schimmack et al., 2017) are particularly social in nature. However, all 29 of them used a fully between-subjects design.

### Illustrative Demonstrations

That between-subjects designs offer less statistical power than within-subjects designs is a fact that is taught in introductory research methods classes. One might think that the

role of this factor in the reproducibility of priming effects would be readily recognized. That has not been the case, from Kahneman’s initial failure to recognize this feature to recent commentary. Shanks (2017) argued that consideration of the within- versus between-subject nature of different priming studies “sheds minimal light on the priming controversy” (p. 1221).

To make the influence of design type on priming results more concrete, we conducted some simple illustrative studies. We selected two robust priming paradigms that employ within-subjects designs from both the cognitive and social psychology literatures, and showed that those robust effects were much more difficult to obtain and yielded less precise effect size estimates in a between- than within-subjects analysis. From the cognitive literature, we chose the Lexical Decision Task (LDT; Meyer & Schvaneveldt, 1971) and the Stroop Task (Stroop, 1935).<sup>3</sup> From the social literature, we chose the Weapons Identification Task (WIT; Payne, 2001) and the Stereotype Misperception Task (SMT; Krieglmeier & Sherman, 2012). For each effect, we analyzed full data sets in two different ways. First, we conducted the standard within-subjects analysis, based on participants’ responses on all trial types (and, thus, all prime-target combinations). In the second approach, each participant was assigned to one level of the experimental design (and, thus, one prime-target combination) based on the first experimental trial to which they had been randomly exposed. In essence, this analysis treated the task as a single trial between-subjects design (for details, see Rivers & Sherman, 2018).

The results were clear. As expected, in the full within-subjects analyses, all four tasks replicated previous work, demonstrating robust and statistically significant priming effects. In contrast, in the between-subjects analysis, though all results were in the expected direction, only the WIT attained (barely) traditional statistical significance ( $p = .048$ ).

The confidence intervals for the effect sizes also differed substantially across the analyses. The average span of 95% confidence intervals for the Hedges  $g$  effect sizes obtained using the within- and between-subjects analyses respectively were .22 versus .75. Simply put, within-subjects designs yielded confidence intervals that spanned a range that was one-third of the size of those from between-subjects designs.

Based on the effect size estimates from the full within-subjects analyses, we estimated how many participants would be needed to achieve power = .8 in both within- and between-subjects designs. For the LDT ( $d_z = .649$ ), a within-subjects design would require 17 participants and a between-subjects design ( $g_s = .203$ ) would require 602 participants. For the Stroop Task ( $d_z = .565$ ), a within-subjects

design would require 21 participants and a between-subjects design ( $g_s = .247$ ) would require 404 participants. For the WIT ( $d_z = .309$ ), a within-subjects design would require 67 participants and a between-subjects design ( $g_s = .496$ ) would require 102 participants. Finally, for the SMT ( $d_z = .292$ ), a within-subjects design would require 74 participants and a between-subjects design ( $g_s = .177$ ) would require 792 participants.

Rudimentary facts about the fixed relationship between design type and power apply in the context of priming research just as they do for other research domains. With equal sample sizes, between-subjects designs yield relatively noisy effect size estimates, are less likely to detect true effects, and are more likely to produce positive results of dubious value.

### What We are not Saying

Here, we address some potential sources of misunderstanding about exactly what we are and are not claiming. First, we are not claiming that between-subjects designs necessarily have less power than within-subjects designs. Of course, it is possible to adequately power between-subjects designs (e.g., by collecting additional observations). The simple reality is that, at the time of the publication of the most noteworthy social priming studies, sample size norms often left between-subjects designs under-powered. All else being equal, the effects demonstrated in those studies would be subsequently less likely to replicate than priming studies published with within-subjects designs.

One might fairly wonder why we focus on the variable of design type rather than power, per se, as it is the direct cause of effect size mis-estimation and publication of false positives. The answer is that it is a simple matter to classify priming studies in terms of research design. It is not such a simple matter to classify them in terms of power. We know that social priming studies differ from other priming studies in their almost exclusive use of between-subjects designs. We cannot definitively state that power covaries perfectly with the distinction between social and nonsocial priming. It is safe to assume that, in general, between-subjects priming studies were less likely to be adequately powered than within-subjects priming studies prior to the “crisis of confidence,” which moved the power issue to the fore. And, in many cases, it can be shown that between-subjects priming effects would have had to assume unrealistically large effect sizes to have been adequately powered (e.g., Bargh et al., 1996).

The results of our demonstration studies do not show that failures to replicate social priming effects can be solely attributed to research design, and we make no such claim. Likewise, we make no claim that social priming effects would all replicate in appropriately powered studies, whether between- or within-subjects. There are several examples of research teams carrying out high-powered replications of between-subjects social priming effects that failed to find evidence for the effects. For example, Gomes and McCullough (2015) were unable to replicate the effect of

<sup>3</sup>We recognize that not everyone considers the Stroop effect to be a priming effect. We believe that it meets the definition in that it captures the unintended influence of a stimulus on associated behavior (i.e., the word and its ink color). Some have argued that the Stroop Task is not a priming task because there is no gap between prime (word meaning) and target (color naming). That is, there is an SOA of 0. However, other priming effects are sometimes implemented with an SOA of 0 (e.g., Gawronski & Ye, 2014; Musch & Klauer, 1997). As well, researchers have implemented SOAs greater than zero (essentially turning Stroop into a sequential priming task) to examine Stroop effects (Glaser & Glaser, 1982; Logan, 1980). In any case, whether you agree that Stroop is priming or not, the methodological point is unchanged.

religious priming on decisions in an economic game despite running 455 participants across two critical between-subjects conditions. Similarly, Shanks et al. (2015), across 9 separate experiments ( $N=1,325$ ), failed to find evidence consistent with previous work showing that priming mating motives affects people's spending behavior. Shanks et al. (2015) specifically had at least 80% power to detect a between-subjects effect as small as  $d_s=.14$ . This highly powered failure to replicate in tandem with demonstrable evidence of publication bias indicates that the reported effect—priming mating motives influences consumer behavior—is most likely a Type I error.

We do not assume that social priming effects that have not been tested with appropriately powered designs are “real” effects. Absent such tests, the existing data are simply unsettled. Even in the case of Bargh et al. (1996), the data are not conclusive. If we assume that the effect size of Bargh et al. (1996) elderly study is comparable to a simple average of the four within-subject priming effect sizes we observed in our illustrative research (a very optimistic estimate in our view), then one would require at least 388 participants over 2 between-subjects conditions to achieve the recommended .8 level of power. If we combine all reported data available on Curate Science (Bargh et al., 1996; Cesario, Plaks, & Higgins, 2006; Doyen et al., 2012; Hull, Slone, Meteyer, & Matthews, 2002; Pashler, Harris, & Coburn, 2008), the resulting sample ( $N=447$ ) would provide the recommended level of power for a single experimental test of the hypothesis. Given the available evidence from the initially reported effect and subsequent failures to replicate, it seems that, at best, the effect must be substantially smaller than initially reported. Still, the exact status of this particular finding is not yet entirely clear.

Finally, we want to be clear that we are not suggesting that statistical power should be prioritized above developing psychological theory or above the study of behavior in natural contexts. There are a multitude of important theoretical questions that cannot be answered using within-subjects approaches and, similarly, there are many difficulties that preclude the use of within-subjects approaches in naturalistic experiments. Many social issues of critical importance are likely only possible to study using between-subjects designs and analyses. We do not wish to suggest that psychologists abandon the study of important psychological principles or of naturalistic behavior simply because within-subjects approaches are not possible. Instead, researchers employing between-subjects designs should seek to maximize statistical power through all means available (e.g., Chartier et al., 2017; Judd, Westfall, & Kenny, 2017; Wang, Sparks, Gonzales, Hess, & Ledgerwood, 2017).

## The Fallout

Had Kahneman and subsequent critics concluded that underpowered published research is unreliable, regardless of its field of origin, very few people would have noticed or cared. However, in associating such work with social psychology and in carelessly tying the work to questions about

research integrity, a firestorm ensued. He framed his letter as an attempt to protect young researchers from the harm of being associated with a “controversial and suspicious” field. In truth, the letter did much more damage than any failed replication. Social priming researchers and social psychologists, more broadly, were hardly alone in paying insufficient attention to statistical power in those days. It had been a long-standing epidemic across much of scientific psychology (see Cohen, 1962; Gigerenzer, 2018). While failing to recognize this fundamental problem, Kahneman saddled the entire field with doubts about integrity. Administrations, departments, and individual scientists who knew little or nothing about social psychology, not to mention priming research in social psychology, suddenly had strong doubts about the field and its researchers. It is difficult to quantify the damage, but few would argue that the episode didn't impact publishing, grant funding, and employment for social psychologists.

Perhaps Kahneman's initial oversight of research design issues was innocent enough, but it is difficult to extend the same good will to some of the subsequent critiques. Responses to Kahneman's letter had brought the design issue to the attention of interested parties, many of whom proceeded to studiously ignore it.<sup>4</sup> It is clear that some of what followed reflected the settling of long-standing grievances about the attention paid to some of the social priming effects (e.g., Dominus, 2017; Zwaan, 2013; see Funder, 2020 for a candid discussion). The *schadenfreude* following replication failures of the research was (and is) thick on social media and in conference hotel bars.

There are a number of ironies in all of this. The social priming work that has received the most attention reports effects that are eye-catching and counter-intuitive. It is the kind of sexy research that popular science writers love to describe, and they were increasingly doing so in the years preceding 2011. Likely for the same reasons, these are the studies that Kahneman (2011) chose to share with his readers in his chapter about priming. The work was attaining great notoriety. This led to some professional resentment, particularly among researchers who felt that there was an unhealthy focus on flashy effects to the detriment of theoretical acuity. One ironic aspect of this is that the same concerns were every bit as prevalent among social psychologists as they were in other fields. There was no shortage of researchers who questioned the reliability of the findings and frowned upon the attention given to those effects. However, the experience of these skeptics changed dramatically when Kahneman threw the whole field of social psychology under the bus, not only with respect to reproducibility, but also with insinuations of outright fraud.

Another ironic aspect of the situation is that the criticism only increased the focus on the flashy findings at the center of the controversy. Tremendous amounts of effort, ink, and bile were spilled fighting about the reproducibility of a small handful of effects that were not and are still not

<sup>4</sup>Kahneman eventually came to recognize the problem as one of statistical power (2017). However, he did not correct the public record with respect to the “social” component of “social priming.”

representative of the research (priming or otherwise) being conducted by social psychologists. In the grand scheme of things, whether those specific effects replicate or not is not of great consequence to the development of psychological science. In the midst of the publication of and furor over these studies, a number of researchers were busy developing broad theoretical models to account for when and how priming works (e.g., Cesario, Plaks, Hagiwara, Navarrete, & Higgins, 2010; Loersch & Payne, 2011; Schröder & Thagard, 2013; Wheeler, DeMarree, & Petty, 2007). This was exactly the sort of work that critics of social priming, both within and outside of social psychology, believed should be getting more attention. However, the impact of this work was blunted as it was crowded out by controversy over a small handful of effects.

## The Future

### *Moderation of Priming*

Moving forward, we concur with many others that the focus should shift from effect-driven debates to theory building and testing (e.g., Doyen, Klein, Simons, & Cleeremans, 2014; Higgins & Eitam, 2014; Molden, 2014). One aspect of such an effort would be to search for important moderators of priming effects. There are many differences among priming studies that may influence their reproducibility. We are not aware of systematic attempts to identify and characterize these moderators. Certainly, no one has systematically studied the impact of content type or researcher home field (e.g., social vs. nonsocial) while controlling for other variables, such as statistical power. We see no a priori theoretical basis for predicting such moderation effects. In contrast, length of delay between prime and target is a potentially important moderator that can be readily examined. Every theory of priming predicts that the effects weaken as the gap between prime and target increases. However, we are not aware of systematic attempts to specify these relationships (for exceptions, see Hermans, De Houwer, & Eelen, 2001; Rivers & Sherman, 2020).

Dividing studies into behavioral and non-behavioral research is conceptually problematic, as described above. To begin, someone would need to develop a coherent and empirically validated taxonomy of types of behavior. One possibly important distinction concerns the extent of constraint on the measured behavior. The robust priming effects typically examined in within-subjects design largely involve two-alternative forced choice tasks. For example, words must be judged as new or old in memory tasks; strings of letters must be judged as words or not in the LDT; words must be evaluated as good or bad in evaluative priming; objects must be judged as a gun or tool in the WIT, and so forth. In contrast, the measures in studies identified as social priming are not constrained to two options. Walking speed, for example, is a continuous behavior. Trait judgments along scales offer more than two options (e.g., Srull & Wyer, 1979). Even performance on a multiple choice knowledge test offers more than two responses (e.g., Dijksterhuis & van Knippenberg, 1998). There certainly are reliable priming

effects with more than two behavioral options (e.g., Coles, Larsen, & Lench, 2019; Strack et al., 1988), but it may be that the robustness of priming effects is related to the constraints on the measured behavior. If the behavior requires a choice between only two options, the influence of a prime may be sufficient to tip the behavior one way or the other. In contrast, with less constrained behavior, the influence of a prime may be insufficient to guide the behavior among multiple or continuous options.

Alternatively, one could make a sensible prediction of exactly the opposite relationship between priming strength and behavioral constraint. Specifically, one might argue that the more constrained the behavior, the less room there is for the behavior to be pushed around by a prime. Deciding whether a stimulus is a word/non-word, positive/negative, etc. is easy. Generally, few errors are made on two-alternative forced choice tasks. If a stimulus is obviously a word or is obviously negative, there may not be sufficient opportunity for the influence of a prime. However, less constrained behaviors, such as walking or making judgments along a 7-point scale, may be more open to external influence.

We don't know if behavioral constraint will ultimately prove to be an important determinant of priming, but it represents the kind of moderator that should be the focus of future research in that it offers a theoretical basis for expecting moderation. Moderators that are post-hoc or that are merely descriptive offer little in the way of theoretical advance.

### *Mechanisms of Priming*

Future research should also focus on the mechanisms that produce priming effects. This has been the primary aim of recent theoretical work (e.g., Cesario et al., 2010; Loersch & Payne, 2011; Schröder & Thagard, 2013; Wheeler et al., 2007). One increasingly common means of developing and testing theoretical models of priming is via the use of formal mathematical models of behavior. Many such models have been proposed to identify and quantify the processes that account for priming effects on memory and judgment. These models also specify the ways in which the component processes interact and constrain one another in producing outcomes. In this way, the models are, in fact, specified theories of the mechanisms underlying priming phenomena. The most commonly used types of models in social (and nonsocial) psychological research are signal detection (e.g., Correll et al., 2002; Green & Swets, 1966), process dissociation (e.g., Jacoby, 1991; Payne, 2001), multinomial processing trees (e.g., Krieglmeier & Sherman, 2012; Meissner & Rothermund, 2013; Payne, Hall, Cameron, & Bishara, 2010; Riefer & Batchelder, 1988; Sherman et al., 2008), and drift diffusion models (e.g., Klauer, Voss, Schmitz, & Teige-Mocigemba, 2007; Pleskac, Cesario, & Johnson, 2018; Ratcliff, 1978).

A comprehensive review of the use of mathematical models to study priming is beyond the scope of this article (see Calanchini, Rivers, Klauer, & Sherman, 2018; Sherman, Klauer, & Allen, 2010), but we do wish to offer an example

from our own research. Recently, we have been using multinomial models to investigate both the moderators and mechanisms of priming effects. This particular work has been conducted in the context of the Stereotype Misperception Task (SMT; Krieglmeyer & Sherman, 2012). In this task, participants are asked to form impressions of people whose faces are shown in blurred drawings. They are asked to judge the faces as high or low on a chosen target trait. Immediately preceding the presentation of the target faces, participants are exposed to prime faces. They are instructed to pay attention to the prime faces, but not respond to them. Any combination of prime types and target traits may be paired. In our studies, the primes are pictures of Black or White males and targets are to be judged as high or low in threateningness. Though the target faces are blurry and ambiguous, there are, in fact, objectively high and low threat versions of the same targets. As such, on each trial there is a correct and incorrect response. The standard SMT priming effect is that the targets are more likely to be judged as threatening when they are preceded by Black face primes than White face primes.

As part of the initial development and validation of the task, Krieglmeyer and Sherman (2012) developed a multinomial model to account for performance and to measure the processes that contribute to it. Without going into great detail, the model proposes that the primes may or may not activate stereotypes that may or may not be corrected in judging the targets. When stereotypes are not activated, responses are determined by the ability to detect the correct responses or response bias. When the activated stereotype (i.e., Black men are threatening) is inconsistent with judging a non-threatening target, the conflicting responses must be resolved.

One of the interesting aspects of the SMT is that it bears close resemblance to the Affect Misattribution Procedure (AMP) developed by Payne et al. (2005), in which participants judge the pleasantness of Chinese pictographs following the presentation of Black and White prime faces. Payne et al. proposed that the AMP effect occurs when positive/negative evaluations of the prime pictures are misattributed to feelings about pictographs and developed a multinomial model to measure this process (Payne et al., 2010). Thus, despite the structural similarities between the SMT and AMP, responses on the tasks are proposed to arise from different processes; specifically, response conflict versus misattribution. And, in fact, our own research has shown that the SMT's response-conflict model reliably offers a better account of SMT performance, whereas the AMP's misattribution model reliably offers a better account of AMP performance. There are a number of differences between the tasks, and why different processes seem to account for performance on the two tasks will be an interesting question for future research.

In the meantime, we wondered if observed priming could be driven by different processes depending on variations in the implementation of the priming task. Specifically, Loersch and Payne (2011) argued that any variable that increases confusion between reactions to a prime and reactions to a

target should increase the likelihood that reactions to the prime will be misattributed as coming from the target. For example, the shorter the delay between prime and target, the more likely a respondent is to confuse reactions to the prime as having come from the target. Source misattribution should also increase to the extent that the signal coming from the target is ambiguous. The more ambiguous the target, the more readily reactions to it may be affected by misattributed reactions to the prime. In a series of studies, we tested both of these potential moderators of priming.

In one study, we manipulated the time that separated the offset of prime images and the onset of target images. We developed three different conditions: The target appeared directly after prime-offset, 50 ms after prime-offset, or 175 ms after prime-offset. We replicated the basic priming effect in all cases. Model fitting showed that the SMT response conflict model and the AMP misattribution model were indistinguishable in terms of model fit indices when prime and target appeared with no time separation. However, with each increase in separation, the advantage of the SMT model increased. At both 50 and 175 ms delay, the SMT model provided significantly better fit to the data than the AMP model. Thus, as prime-target separation increased, so too did the extent to which the response conflict process model better accounted for priming than did the misattribution process model. If the mechanisms underlying priming effects shift within a few hundred milliseconds, it is a safe bet that the longer gaps frequently used in social priming studies will also involve a variety of different psychological mechanisms.

In another study, we manipulated the level of ambiguity in the target faces. In one condition, the high and low threat targets were 3 SDs above and below a mean level of threat. In the other condition, the high and low threat targets were, in fact, equal in objective level of threat. Once again, we replicated the basic priming effect both when targets were unambiguous and when targets were fully ambiguous. Whereas the SMT process model provided significantly better fit when the targets were unambiguous, the fits of the two models were indistinguishable when the targets were ambiguous. Thus, increasing target ambiguity reduced the extent to which response conflict processes better accounted for priming than misattribution.

We provide these as examples of the kinds of questions about priming moderation and mechanism that we view as useful for developing a deeper theoretical understanding of when, how, and why priming effects occur. Modeling is certainly not the only way to conduct such investigations, and other strategies may be equally or more effective.

## Conclusion

Does it matter how we characterize the research that has been labeled "social priming?" We think that it does. First and foremost, if we want to improve our research practices so that our work may be more reliable, we need to accurately identify the reasons why research is more or less robust. To the extent that we misidentify and misunderstand

the causes of unreliable research, scientific progress will slow. In learning the wrong lessons, we also may fail to learn the right ones.

Beyond the issue of accuracy, how we account for variations in the reproducibility of research has real implications for real people and the work that they do. As a rule, we would suggest that commenters should be cautious about ascribing specific systematic causes for replication failure in the absence of thorough empirical examination. Certainly, unsubstantiated insinuations of unethical behavior are uncalled for, particularly when there are more mundane explanations for failures to replicate.

As it stands, the term "social priming" offers nothing in the way of explaining the relative robustness of different priming effects. Moreover, use of that term has obscured important features that do help to account for the relative robustness of priming effects. The studies discussed and examined as cases of social priming are best described as priming studies that employ between-subjects designs. The content of the primes and targets is not necessarily social. As well, priming research with clearly social content is not counted as social priming when it uses within-subjects designs, even if the work is conducted by social psychologists. There are reasons why social psychological priming research may be more likely to rely on between-subjects designs than priming research in other fields. In particular, concerns about carryover effects in within-subjects designs may be relatively more common in social psychological research (see Greenwald, 1976). Nevertheless, a large majority of priming research in social psychology has used within-subjects designs that have yielded reliable results.

It is well past time to finally put to rest the label "social priming." It also is past time to move beyond debates about specific findings. We should have more ambitious goals than merely testing the reliability of particular effects. We should aim for a broader theoretical understanding of the moderators and mechanisms underlying priming in order to better predict when, why, and how priming occurs.

## Acknowledgement

Preparation of this manuscript was partially supported by an Anneliese Maier Research Award from the Alexander von Humboldt Foundation to JWS.

## References

- Bargh, J. A. (2005). Bypassing the will: Toward demystifying the non-conscious control of social behavior. In R. R. Hassin, J. S. Uleman, & J. A. Bargh (Eds.), *The new unconscious* (pp. 37–60). Oxford: Oxford University Press.
- Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype-activation on action. *Journal of Personality and Social Psychology*, 71(2), 230–244. doi:10.1037//0022-3514.71.2.230
- Becker, S., Moscovitch, M., Behrmann, M., & Joordens, S. (1997). Long-term semantic priming: A computational account and empirical evidence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(5), 1059–1082.
- Bem, D. J. (2011). Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology*, 100(3), 407–425. doi:10.1037/a0021524
- Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 14(5), 365–376. doi:10.1038/nrn3475
- Calanchini, J., Rivers, A. M., Klauer, K. C., & Sherman, J. W. (2018). Multinomial processing trees as theoretical bridges between cognitive and social psychology. *Psychology of Learning and Motivation*, 69, 39–65. doi: <http://doi.org/10.1016/bs.plm.2018.09.002>
- Cameron, C. D., Brown-Iannuzzi, J. L., & Payne, B. K. (2012). Sequential priming measures of implicit social cognition: A meta-analysis of associations with behavior and explicit attitudes. *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc.*, 16(4), 330–350.
- Cesario, J., Plaks, J. E., Hagiwara, N., Navarrete, C. D., & Higgins, E. T. (2010). The Ecology of automaticity. How situational contingencies shape action semantics and social behavior. *Psychological Science*, 21(9), 1311–1317. doi:10.1177/0956797610378685
- Cesario, J., Plaks, J. E., & Higgins, E. T. (2006). Automatic social behavior as motivated preparation to interact. *Journal of Personality and Social Psychology*, 90(6), 893–910. doi:10.1037/0022-3514.90.6.893
- Chartier, C. R., McCarthy, R. J., Williams, S. R., Ebersole, C. R., Hamlin, K., Lucas, R. E., ... Lenne, R. L. (2017). StudySwap: A platform for interlab replication, collaboration, and research resource exchanged. Retrieved from [osf.io/9aj5g](https://osf.io/9aj5g)
- Chivers, T. (2019). What's next for psychology's embattled field of social priming. *Nature*, 576(7786), 200–202. doi:10.1038/d41586-019-03755-2
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum.
- Cohen, J. (1962). The statistical power of abnormal-social psychological research: A review. *Journal of Abnormal and Social Psychology*, 65, 145–153. doi:10.1037/h0045186
- Coles, N. A., Larsen, J. T., & Lench, H. C. (2019). A meta-analysis of the facial feedback literature: Effects of facial feedback on emotional experience are small and variable. *Psychological Bulletin*, 145(6), 610–651. doi:10.1037/bul0000194
- Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2002). The police officer's dilemma: Using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality and Social Psychology*, 83(6), 1314–1329.
- Dijksterhuis, A., & van Knippenberg, A. (1998). The relation between perception and behavior, or how to win at a game of Trivial Pursuit. *Journal of Personality and Social Psychology*, 74(4), 865–877. [Database] doi:10.1037/0022-3514.74.4.865
- Dominus, S. (2017, October). When the revolution came for Amy Cuddy. *The New York Times Magazine*.
- Doyen, S., Klein, O., Pichon, C., & Cleeremans, A. (2012). Behavioral priming: It's all in the mind, but whose mind? *PLoS One*, 7(1), e29081. doi:10.1371/journal.pone.0029081
- Doyen, S., Klein, O., Simons, D. J., & Cleeremans, A. (2014). On the other side of the mirror: Priming in cognitive and social psychology. *Social Cognition*, 32(Supplement), 12–32. doi:10.1521/soco.2014.32.sup.12
- Engber, D. (2017, June 7). Daryl Bem proved ESP is real: Which means science is broken. *Slate*. Retrieved from <https://slate.com/health-and-science/2017/06/daryl-bem-proved-esp-is-real-showed-science-is-broken.html>.
- Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, 69(6), 1013–1027. doi:10.1037//0022-3514.69.6.1013
- Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, 50(2), 229–238. doi:10.1037/0022-3514.50.2.229

- Funder, D. C. (2020, May 8). Thoughts on “Ego Depletion” and some related issues concerning replication. *Funderstorms*, Retrieved from <https://funderstorms.wordpress.com/2020/05/08/thoughts-on-ego-depletion-and-some-related-issues-concerning-replication/>.
- Gawronski, B., & Ye, Y. (2014). What drives priming effects in the affect misattribution procedure? *Personality & Social Psychology Bulletin*, 40(1), 3–15. doi:10.1177/0146167213502548
- Gigerenzer, G. (2018). Statistical rituals: The replication delusion and how we got there. *Advances in Methods and Practices in Psychological Science*, 1(2), 198–218. doi:10.1177/2515245918771329
- Gilder, T. S. E., & Heerey, E. A. (2018). The role of experimenter belief in social priming. *Psychological Science*, 29(3), 403–417. doi:10.1177/0956797617737128
- Glaser, M. O., & Glaser, W. R. (1982). Time course analysis of the Stroop phenomenon. *Journal of Experimental Psychology: Human Perception and Performance*, 8(6), 875–894. doi:10.1037//0096-1523.8.6.875
- Gomes, C. M., & McCullough, M. E. (2015). The effects of implicit religious primes on dictator game allocations: A preregistered replication experiment. *Journal of Experimental Psychology: General*, 144(6), e94–104. doi:10.1037/xge0000027
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Greenwald, A. G. (1976). Within-subjects designs: To use or not to use. *Psychological Bulletin*, 83(2), 314–320. doi:10.1037/0033-2909.83.2.314
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6), 1464–1480. doi:10.1037/0022-3514.74.6.1464
- Harris, C. R., Coburn, N., Rohrer, D., & Pashler, H. (2013). Two failures to replicate high-performance-goal priming effects. *PLoS One*, 8(8), e72467. doi:10.1371/journal.pone.0072467
- Hermans, D., De Houwer, J., & Eelen, P. (2001). A time course analysis of the affective priming effect. *Cognition & Emotion*, 15(2), 143–165. doi:10.1080/02699930125768
- Higgins, E. T., Eitam, B. (2014). Priming...shmiming: It's about knowing when and why stimulated memory representations become active. *Social Cognition*, 32(Supplement), 225–242. doi:10.1521/soco.2014.32.sup.225
- Hughes, A. D., & Whittlesea, B. W. A. (2003). Long-term semantic transfer: An overlapping-operations account. *Memory & Cognition*, 31(3), 401–411. doi:10.3758/bf03194398
- Hull, J. G., Slone, L. B., Meteyer, K. B., & Matthews, A. R. (2002). The nonconsciousness of self-consciousness. *Journal of Personality and Social Psychology*, 83(2), 406–424. doi:10.1037//0022-3514.83.2.406
- Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, 30(5), 513–541. doi:10.1016/0749-596X(91)90025-F
- Judd, C. M., Westfall, J., & Kenny, D. A. (2017). Experiments with more than one random factor: Designs, analytic models, and statistical power. *Annual Review of Psychology*, 68, 601–625. doi:10.1146/annurev-psych-122414-033702
- Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus, and Giroux.
- Kahneman, D. (2012). A proposal to deal with questions about priming effects (open letter via email).
- Kahneman, D. (2017, February 14). Re: Reconstruction of a train wreck: How priming research went off the rails [Blog comment]. Retrieved from: <https://replicationindex.com/2017/02/02/reconstruction-of-a-train-wreck-how-priming-research-went-off-the-rails/comment-page-1/#comments>.
- Klauer, K. C., Voss, A., Schmitz, F., & Teige-Mocigemba, S. (2007). Process components of the Implicit Association Test: A diffusion-model analysis. *Journal of Personality and Social Psychology*, 93(3), 353–368. doi:10.1037/0022-3514.93.3.353
- Krieglmeyer, R., & Sherman, J. W. (2012). Disentangling stereotype activation and stereotype application in the Stereotype Misperception Task. *Journal of Personality and Social Psychology*, 103(2), 205–224. doi:10.1037/a0028764
- Levelt Committee, Noort Committee, Drenth Committee. (2012). *Flawed science: The fraudulent research practices of social psychologist Diederik Stapel*. Retrieved from <http://www.tilburguniversity.edu/nl/nieuws-en-agenda/finalreportLevelt.pdf>.
- Loersch, C., & Payne, B. K. (2011). The situated inference model: An integrative account of the effects of primes on perception, behavior, and motivation. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, 6(3), 234–252. doi:10.1177/1745691611406921
- Logan, G. D. (1980). Attention and automaticity in the Stroop and priming tasks: Theory and data. *Cognitive Psychology*, 12(4), 523–553. doi:10.1016/0010-0285(80)90019-5
- McCook, A. (2017, February 20). I placed too much faith in underpowered studies: Nobel Prize winner admits mistakes. *Retraction Watch*.
- Meissner, F., & Rothermund, K. (2013). Estimating the contributions of associations and recoding in the Implicit Association Test: The ReAL model for the IAT. *Journal of Personality and Social Psychology*, 104(1), 45–69. doi:10.1037/a0030734
- Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, 90(2), 227–234. doi:10.1037/h0031564
- Meyer, M. L. (2019). Social by default: Characterizing the social functions of the resting brain. *Current Directions in Psychological Science*, 28(4), 380–386. doi:10.1177/0963721419857759
- Molden, D. (2014). Understanding priming effects in social psychology: What is “social priming” and how does it occur? *Social Cognition*, 32(Supplement), 1–11. doi:10.1521/soco.2014.32.sup.1
- Musch, J., & Klauer, K. C. (1997). The proportion affect in affective priming: Replication and evaluation of a theoretical explanation. *Zeitschrift Experimental Psychology*, 44, 266–292.
- Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. *Basic Processes in Reading: Visual Word Recognition*, 11, 264–336. DOI: 10.3758/BF03212427
- Neuroskeptic. (2017, August 16). More on ‘behavior priming’ and unconscious influences. *Discover Blogs*. Retrieved from <http://blogs.discovermagazine.com/neuroskeptic/2017/08/16/behavior-priming-controversy/>.
- Neuroskeptic. (2018, January 26). What is “social priming”? *Discover Blogs*. Retrieved from <https://www.discovermagazine.com/the-sciences/what-is-social-priming>.
- Nosek, B. A., & Banaji, M. R. (2001). The Go/No-go association task. *Social Cognition*, 19(6), 625–666. [Database] doi:10.1521/soco.19.6.625.20886
- Pashler, H., Harris, C., & Coburn, N. (2008). Elderly-related words prime slow walking. Retrieved from <http://www.PsychFileDrawer.org/replication.php?attempt=MTU%3D>.
- Pashler, H., Rohrer, D., & Harris, C. (2013). Can the goal of honesty be primed? *Journal of Experimental Social Psychology*, 14, 959–964. DOI: 10.1016/j.jesp.2013.05.011
- Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology*, 81(2), 181–192. doi:10.1037/0022-3514.81.2.181
- Payne, B. K., Brown-Iannuzzi, J. L., & Loersch, C. (2016). Replicable effects of primes on human behavior. *Journal of Experimental Psychology: General*, 145(10), 1269–1279. doi:10.1037/xge0000201
- Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology*, 89(3), 277–293.
- Payne, B. K., Hall, D., Cameron, C. D., & Bishara, A. J. (2010). A process model of affect misattribution. *Personality & Social Psychology Bulletin*, 36(10), 1397–1408. doi:10.1177/0146167210383440
- Pleskac, T. J., Cesario, J., & Johnson, D. J. (2018). How race affects evidence accumulation during the decision to shoot. *Psychonomic Bulletin & Review*, 25(4), 1301–1330. doi:10.3758/s13423-017-1369-6
- Prabhakaran, R., & Gray, J. R. (2012). The pervasive nature of unconscious social information processing in executive control. *Frontiers in Human Neuroscience*, 6, 105. doi:10.3389/fnhum.2012.00105

- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85(2), 59–108. doi:10.1037/0033-295X.85.2.59
- Riefer, D. M., & Batchelder, W. H. (1988). Multinomial modeling and the measurement of cognitive processes. *Psychological Review*, 95(3), 318–339. doi:10.1037/0033-295X.95.3.318
- Rivers, A. M., & Sherman, J. W. (2018, January 19). Experimental design and the reliability of priming effects: Reconsidering the "train wreck". doi:10.31234/osf.io/r7pd3
- Rivers, A. M., & Sherman, J. W. (2020). *Same task, different processes: Different implementations of an implicit measure invoke differing underlying mechanisms of bias*. Unpublished manuscript.
- Schimmack, U., Heene, M., & Kesavan, K. (2017, February 2). Reconstruction of a train wreck: How priming research went off the rails. *Replicability Index*. Retrieved from <https://replicationindex.wordpress.com/2017/02/02/reconstruction-of-a-train-wreck-how-priming-research-went-off-the-rails>.
- Schröder, T., & Thagard, P. (2013). Priming: Constraint satisfaction and interactive competition. *Social Cognition*, 32(Supplement), 152–172. doi:10.1521/soco.2014.32.supp.152
- Shanks, D. R. (2017). Misunderstanding the behavior priming controversy: Comment on Payne, Brown-Iannuzzi, and Loersch (2016). *Journal of Experimental Psychology: General*, 146(8), 1216–1222. doi:10.1037/xge0000307
- Shanks, D. R., Newell, B. R., Lee, E. H., Balakrishnan, D., Ekelund, L., Cenac, Z., ... Moore, C. (2013). Priming intelligent behavior: An elusive phenomenon. *PLoS One*, 8(4), e56515. <https://doi.org/10.1371/journal.pone.0056515> doi:10.1371/journal.pone.0056515
- Shanks, D. R., Vadillo, M. A., Riedel, B., Clymo, A., Govind, S., Hickin, N., ... Puhlmann, L. M. C. (2015). Romance, risk, and replication: Can consumer choices and risk-taking be primed by mating motives. *Journal of Experimental Psychology: General*, 144(6), e142–158. doi:10.1037/xge0000116
- Sherman, J. W., Gawronski, B., Gonsalkorale, K., Hugenberg, K., Allen, T. J., & Groom, C. J. (2008). The self-regulation of automatic associations and behavioral impulses. *Psychological Review*, 115(2), 314–335. doi:10.1037/0033-295X.115.2.314
- Sherman, J. W., Klauer, K. C., & Allen, T. J. (2010). Mathematical modeling of implicit social cognition: The machine in the ghost. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 156–175). New York: Guilford Press.
- Sherman, J. W., & Rivers, A. M. (2020). Social priming: time to ditch a dubious term? *Nature*, 579(7797), 29. doi:10.1038/d41586-020-00616-1
- Sherman, S. M., & Jordan, T. R. (2011). Word-frequency effects in long-term semantic priming and false memory. *British Journal of Psychology*, 102(3), 559–568. doi:10.1111/j.2044-8295.2011.02017.x
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22(11), 1359–1366. doi:10.1177/0956797611417632
- Spellman, B. A. (2015). A short (personal) future history of revolution 2.0. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, 10(6), 886–899. doi:10.1177/1745691615609918
- Srull, T. K., & Wyer, R. S. (1979). The role of category accessibility in the interpretation of information about persons: Some determinants and implications. *Journal of Personality and Social Psychology*, 37(10), 1660–1672. doi:10.1037/0022-3514.37.10.1660
- Strack, F., Martin, L. L., & Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile: A nonobtrusive test of the facial feedback hypothesis. *Journal of Personality and Social Psychology*, 54(5), 768–777. doi:10.1037//0022-3514.54.5.768
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18(6), 643–662. doi:10.1037/h0054651
- Trope, Y., & Fishbach, A. (2005). Going beyond the motivation given: Self-control and situational control over behavior. In R. R. Hassin, J. S. Uleman, & J. A. Bargh (Eds.), *The new unconscious* (pp. 537–566). Oxford: Oxford University Press.
- Tse, C.-S., & Neely, J. H. (2005). Assessing activation without source monitoring in the DRM false memory paradigm. *Journal of Memory and Language*, 53(4), 532–550. doi:10.1016/j.jml.2005.07.001
- Tse, C.-S., & Neely, J. H. (2007). Semantic priming from letter-searched primes occurs for low- but not high-frequency targets: Automatic semantic access may not be a myth. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 1143–1161. doi:10.1037/0278-7393.33.6.1143
- Wagenmakers, E.-J. (2014, March 26). Behavioral priming: Time to nut up or shut up. Open Science Collaboration. Retrieved from <http://osc.centerforopenscience.org/2014/03/26/behavioral-priming/>.
- Wang, Y. A., Sparks, J., Gonzales, J. E., Hess, Y. D., & Ledgerwood, A. (2017). Using independent covariates in experimental designs: Quantifying the trade-off between power boost and Type I error inflation. *Journal of Experimental Social Psychology*, 72, 118–124. doi:10.1016/j.jesp.2017.04.011
- Weingarten, E., Chen, Q., McAdams, M., Yi, J., Hepler, J., & Albarracín, D. (2016). From primed concepts to action: A meta-analysis of the behavioral effects of incidentally presented words. *Psychological Bulletin*, 142(5), 472–497. doi:10.1037/bul0000030
- Wentura, D., & Rothermund, K. (2014). Priming is not priming is not priming. *Social Cognition*, 32(Supplement), 47–67. doi:10.1521/soco.2014.32.supp.47
- Wheeler, S. C., DeMarree, K. G., & Petty, R. E. (2007). Understanding the role of the self in prime-to-behavior effects: The active-self account. *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc*, 11(3), 234–261. doi:10.1177/1088868307302223
- Woltz, D. J., & Was, C. A. (2006). Availability of related long-term memory during and after attention focus in working memory. *Memory & Cognition*, 34(3), 668–684. doi:10.3758/bf03193587
- Woltz, D. J., & Was, C. A. (2007). Available but unattended conceptual information in working memory: Temporarily active semantic content or persistent memory for prior operations? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(1), 155–168. doi:10.1037/0278-7393.33.1.155
- Yong, E. (2012, March 10). A failed replication draws a scathing personal attack from a psychology professor. Discover blogs: Not Exactly Rocket Science. Retrieved from <http://blogs.discovermagazine.com/notrocketscience/2012/03/10/failed-replication-bargh-psychology-study-doyen/>.
- Zhong, C.-B., & Liljenquist, K. (2006). Washing away your sins: Threatened morality and physical cleansing. *Science*, 313(5792), 1451–1452. doi:10.1126/science.1130726
- Zwaan, R. (2013, September 17). How to cook up your own social priming article. Zeistgeist: Psychological Experimentation, Cognition, Language, and Academia. Retrieved from <https://rolfzwaan.blogspot.com/2013/09/how-to-cook-up-your-own-social-priming.html>.