Priming, Replication, and the Hardest Science

Joseph Cesario

Perspectives on Psychological Science 2014 9: 40
DOI: 10.1177/1745691613513470

aps ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Perspectives on Psychological Science
2014, Vol 9(1) 40–48

DOI: 10.1177/1745691613513470
pps.sagepub.com
SAGE

Priming, Replication, and the Hardest Science

Joseph Cesario
Michigan State University

Abstract

Concerns have been raised recently about the replicability of behavioral priming effects, and calls have been issued to identify priming methodologies with effects that can be obtained in any context and with any population. I argue that such expectations are misguided and inconsistent with evolutionary understandings of the brain as a computational organ. Rather, we should expect priming effects to be highly sensitive to variations in experimental features and subject populations. Such variation does not make priming effects frivolous or capricious but instead can be predicted *a priori*. However, absent theories specifying the precise contingencies that lead to such variation, failures to replicate another researcher's findings will necessarily be ambiguous with respect to the inferences that can be made. Priming research is not yet at the stage where such theories exist, and therefore failures are uninformative at the current time. Ultimately, priming researchers themselves must provide direct replications of their own effects; researchers have been deficient in meeting this responsibility and have contributed to the current state of confusion. The recommendations issued in this article reflect concerns both with the practice of priming researchers and with the inappropriate expectations of researchers who have failed to replicate others' priming effects.

Keywords

priming, replication, automatic social behavior, prime to behavior, goal priming

Editor's Note: Behavioral priming has come under intense scrutiny during the recent methodological self-questioning in psychological science. Major concerns have been raised about the replicability and robustness of the findings. Last year, *Perspectives* received three independent articles that speak to the issues of what should be expected from attempted replications and the value of attempted replications, both for behavioral priming and in general. Those articles are grouped together here along with two commentaries solicited from leading researchers with very different ideas about the lessons to be drawn from them.

Recent years have seen a rapidly growing interest in replication within psychology and concern over failures to replicate published findings, as seen in the many publications on this topic (e.g., Pashler & Wagenmakers, 2012; Spellman, 2013), the creation of a special registered replication section in *Perspectives on Psychological Science*, and the increasing activity on Web sites such as PsychFileDrawer.org. Replication can accomplish a number of important goals, such as narrowing effect size estimates and providing information about whether an earlier published effect should be considered to be Type I error. Such efforts are absolutely essential if psychology is to develop as a mature science.

Although concerns with replicability have spanned most every area of psychology and beyond (Gelman, in press; Ioannidis, 2005), priming research has been a particular target for those who care about replication in psychology—this was perhaps encouraged by Daniel Kahneman's letter highlighting concerns in this area.[1] Most important, there has been a desire from critics to identify invariant priming effects that can be consistently and readily obtained by any researcher anywhere in the world, on the Internet, in heavy rain, and so forth. When replication attempts fail, many critics have been quick to conclude either that the original demonstrations were Type I error or that priming effects are too fickle to be of interest. Several high profile replication failures of classic priming effects have elicited such reactions, and in discussions with colleagues from other areas, one gets the sense that such failures fuel further disbelief and suspicion about this area of research.

At the outset, it is important to be clear about the purpose of this article. The purpose is not to review

Corresponding Author:
Joseph Cesario, Department of Psychology, Michigan State University, East Lansing, MI 48824
E-mail: cesario@msu.edu

evidence concerning any specific priming effect, nor is it even to instill in the reader confidence in priming effects in general. It might be the case that priming effects are not real and published effects are a combination of pure Type I error and researcher degrees of freedom (Simmons, Nelson, & Simonsohn, 2011). Without question, we must be open to this possibility and we must reevaluate our conclusions as new empirical evidence comes to light (or fails to come to light, if priming researchers do not or cannot replicate their own findings).

Instead, the purpose of this article is to discuss the nature of replication, how it relates to psychological science, and why priming research in its current state may hold particular challenges regarding inferences about replication results. Specifically, what are the appropriate expectations for the replicability of priming results and when are independent replications informative? I argue that we should expect priming effects to vary by a wide range of moderating individual difference and experimental context variables, and absent well-developed theories for specifying such variables, the conclusions of replication failures will be ambiguous. Furthermore, the expectation that priming effects should be widely invariant is inconsistent with what is known about the evolved, computational nature of the mind. Given this variability, one must take the long view and allow this area to develop adequately before premature death knells are sounded.

Ultimately, priming researchers themselves need to provide repeated replications of their own work upon initial publication. Doing so will make it clear that subsequent replication failures by different laboratories are indicative of important contextual moderation worthy of further study and not indicative of Type I error of the original findings.

The Importance of Replication: What Does It Do for Us?

As a starting point, I take as a given that replication is an essential research activity for science. Although nearly every scientist agrees on this general point, few explicitly discuss why this is true. What does replication tell us about the nature of science, and how does it allow us to distinguish science from other means of knowing the world (such as religious approaches)?

Heidbreder (1933) argued that when we say "scientific knowledge is objective," there is a very precise meaning at play for the terms *objective* and *subjective*. Objectivity here does not refer to metaphysical claims about what exists outside one's own mind. When we say that physics yields objective knowledge, we are not making claims about the metaphysical status of the substances that physicists study. In point of fact, it is this disregard for the question of metaphysics that allowed physics and other natural sciences to advance as mature sciences. Instead, what we mean by objective knowledge is that such knowledge has the quality of social verification. As Heidbreder stated:

> It is not the source of knowledge but its verification that determines its objectivity in science. The physicist does not pause at length to ask: Do these events which I observe exist only in my perception? He asks: Do they lead to communicable judgments based on repeated and verified experience? (p. 272)

Replication is important not for its own sake, but because it is a sign that the knowledge being generated will pass the social verification test. We know that physics generates objective knowledge precisely because we will both get the same answer if you measure the rate of the ball dropped from the tower and I measure it as well. I can observe your stopwatch and confirm its reading; we (and others in the field) agree that this is an appropriate measurement device to begin with. It is only through the process of doing the same thing and arriving at the same conclusion that we have a science that generates knowledge we call objective.

What Is Replication?

If replication is the key to objective knowledge, what exactly is replication and how does it play out in the area of priming? If replication is to serve the function of social verification of generated knowledge, then it must mean "doing the same thing" in terms of the methods used in an original investigation. On the surface, this appears quite straightforward, and in some cases it will be. If a researcher finds that responses on two questionnaires positively correlate in a sample of Michigan State University students, it is a mostly straightforward matter to replicate this study—that is, to do the same thing in terms of the methodology. I can distribute in the same way those same questionnaires to another sample recruited in the same way from this same population. If I repeat this sampling 100 times, one of two things will happen: (a) I generate a large proportion of nonzero positive correlations and get a more precise estimate of the true population effect size, or (b) I fail to generate a substantive number of positive correlations and we conclude that the original finding was likely Type I error.[2]

If some feature of the methodology is changed, then the inference one can draw about a set of failed replications becomes ambiguous. That is, being able to infer that the original finding was Type I error depends absolutely on the methodology being the same as in the original finding. If and only if I did the same thing can I conclude that my replication failures indicate Type I error. If I change some important variable of the methodology, then all bets are off with respect to the inference I can draw about that failure. (As explicated in more detail

below, a change in the population from which a researcher is sampling constitutes a change in methodology, in which case having a theory for predicting and explaining such differences becomes of central importance in interpreting a failure to replicate; see also Asendorpf et al., 2013; Henrich, Heine, & Norenzayan, 2010.)

Priming Methodologies and Replication

Critics of priming methodologies not only desire direct replications of priming effects (which is reasonable), they have also voiced the expectation that priming effects should not vary by any but the most obvious and salient situation or population characteristics. When replication failures occur, critics frequently conclude either that original effect was likely Type I error or that the effect in question is so sensitive as to be unimportant and capricious. But are these reasonable conclusions?

The first step in answering this question is to ask whether we should, in fact, expect any but the most basic cognitive and perceptual priming effects to be consistent across broad segments of populations and across variable experimental contexts. I argue that this expectation is entirely misguided and inconsistent with the predominant understanding of the brain as a computational organ that uses informational inputs to execute adaptive decision rules. To illustrate, consider the well-known priming effect obtained by Bargh, Chen, and Burrows (1996) that subliminally priming pictures of Black males results in more aggressive responses to provocation (relative to priming pictures of White males). What are the appropriate expectations for the replicability of this effect?

We know that there are some features of the original methodology that must be reproduced exactly for a replication attempt to be informative because we know that changing them would change the nature of the effect. So, for example, we would want to use exactly the same pictures of Black and White males as in the original, the same prime duration, and the same magnitude of provocation. But how do we know that these features are important and that changing them would change the original effect (thereby making the results of the replication attempt noninformative)? We know this only because we have relevant theories that tell us that these features should matter. With respect to the types of pictures used as primes, for instance, we know that people have different automatic responses to Blacks with high or low Afrocentric features or with dark or light skin (Blair, Judd, & Fallman, 2004; Hagiwara, Kashy, & Cesario, 2012; Ma & Correll, 2011), which would lead to different predicted effects if the pictures varied along these dimensions. Similarly, we know that prime duration influences the strength of a prime, and shorter prime durations might eliminate the effects of priming young Black males (Higgins, 1996).

The central problem with replications in priming research is that, at the current time, we do not yet have enough well-developed formal theories of priming that can pinpoint exactly what features and what effects should be important. Because of this, replication failures at this stage will necessarily be ambiguous because we cannot be sure that features that appear incidental to the researcher are not actually integral to obtaining the original effect. This is not a problem inherent to priming in general, or priming social behaviors, or priming motivation. It is simply a function of the relatively young state of priming research.[3]

To illustrate this point, I have provided a few examples of some moderating variables that have been uncovered since Bargh and colleagues published the priming effect described earlier but that were probably unknown to researchers at the time. These include both individual difference variables and experimental variables; thus, changes to the population being sampled or changes to the experimental context would render replication failures ambiguous (yet still may provide meaningful information about underlying process.)

In terms of variables that might be specific to the effect of priming Black males on aggressive responses, my colleagues and I found that the degree to which an individual stereotypes Black males as dangerous moderates this effect (Cesario, Plaks, Hagiwara, Navarrete, & Higgins, 2010; see also Dijksterhuis, Aarts, Bargh, & van Knippenberg, 2000). This is based on our understanding that aggressive behavioral responses to the priming event in the Bargh et al. paradigm are the output of participants preparing to interact with a dangerous outgroup male (see Cesario, Plaks, & Higgins, 2006). To the extent that participants do not represent Black males as dangerous, then there would be no need to prepare and execute defensive threat behavior. In addition, one might expect that a participant's physical size should also influence responses to the primes in this paradigm. Given that physically larger individuals can perform aggressive behaviors more effectively and incur lower costs doing so than would smaller individuals, it follows that one's physical formidability may influence aggression in response to the provocation (Sell, Tooby, & Cosmides, 2009).

Besides such individual differences, one can also identify features of the experimental context that, if varied, could eliminate or even reverse priming effects. Recently, we found that the physical surroundings in priming experiments could influence participants' responses to primes, given that different physical situations allow for different actions to be executed (Cesario et al., 2010). Drawing on the nonhuman animal literature on defensive

threat behavior, we predicted and found that White participants who were seated in a physical space that restricted their ability to escape (a sound-resistant booth) showed increased accessibility of fight-related action semantics when primed with pictures of Black males. However, when primed with the same pictures while seated in a physical space that allowed distancing behavior (an open field), participants instead showed increased accessibility of flight-related action semantics. Similarly, we have found that priming participants while they are surrounded by ingroup members changes the nature of the response to various primes, given that being around one's coalition allows behaviors to be executed that could not be performed easily by oneself (Cesario & Jonas, 2013; see also Fessler & Holbrook, 2013). Therefore collecting data individually versus in a group setting could eliminate or reverse a sample-wide average response to a prime.

In addition, the outcomes of replication attempts could be influenced by a range of variables now known to influence priming effects in general (not just the specific Black male priming effect described above): interdependent self-construals (rather than independent self-construals) lead to assimilative priming effects (Bry, Follenfant, & Meyer, 2008), low (rather than high) self-monitors are more likely to show priming effects (DeMarree, Wheeler, & Petty, 2005), the positivity associated with a prime influences responses (Custers & Aarts, 2005), associating a primed goal with ingroup or outgroup members can change responses to primes (Loersch, Aarts, Payne, & Jefferis, 2008), and the prime's effect can change based on whether the prime is self- or other-generated (Mussweiler & Neumann, 2000; see also Mussweiler & Förster, 2000).

The point here is that if some critical feature of the experimental context has been changed, even if it is deemed irrelevant by the researcher, we would expect such a change to eliminate or reverse a previous priming effect (see also Katzko, 2006). Similarly, if a researcher is sampling from a population that differs markedly on some key individual difference variable from the population sampled by the original researcher, such as drawing from a population of chronically high or low self-monitors, then we would expect that the priming effect will not be obtained (or will even reverse).

As noted above, given the infancy of priming research we do not yet know all the relevant moderators and process-related variables; thus, there will be a host of variables that we fail to identify as important and that may be changed unwittingly by a replicating researcher. To the extent that, as noted earlier, "doing the same thing" is required to draw accurate inferences about replication failures, then not having a full accounting of the relevant variables will prevent meaningful conclusions when a different researcher fails to replicate a priming effect.[4]

Is the Search for Invariant Effects Even Reasonable?

The argument that priming effects should vary considerably is not post hoc excuse making for failures to replicate. Variability should be expected because this is what available data would suggest with respect to important social and motivated behaviors. Varying responses (across individuals, across contexts) to an identical stimulus is the norm for both human and nonhuman animals. To provide just a few examples illustrating the range of behaviors for which this is true, we know this to be the case for defensive threat behavior in rodents (R. J. Blanchard, Flannelly, & Blanchard, 1986) and humans (D. C. Blanchard, Hynd, Minke, Minemoto, & Blanchard, 2001), hormonal effects on aggression and intimacy (Bos, Panksepp, Bluthé, & van Honk, 2012), effects of chronically active self-concept attributes (Brown & McConnell, 2009), and the effects of physical postures on power (Cesario & McDonald, 2013). With the exception of purely reflexive behaviors, it is difficult to find many examples in which behavior is executed inflexibly in response to a stimulus without regard to other information in the surrounding environment.

The sensitivity of priming effects to a range of individual difference and experimental context variables is consistent with current approaches to understanding the evolved functions of the mind. The widely held perspective from evolutionary psychology is that the mind is a computational organ designed to incorporate information from different sources to regulate behavior. Adaptive decision rules are followed using such information as inputs (see Tooby & Cosmides, 2005).

With respect to social interactions, the brain cannot successfully regulate the body to any reasonable degree if it does not take into account information about the self, target others, and the broader contexts in which interactions take place. In other words, the brain must operate on information beyond target stimulus features. In terms of priming an aggressive outgroup male, decisions about whether to execute costly aggressive behaviors and escalate a conflict involve a number of variables beyond simply the aggressiveness of the target other (i.e., whether the target is stereotyped as aggressive). Such decisions include, but are not limited to, a calculus of (a) one's own physical formidability, (b) the formidability of the opponent, (c) external contingencies such as barriers to escape and distance to the threat, and (d) assessments of one's own coalitional support relative to the coalitional support of the opponent (see, e.g., Benson-Amram,

Heinen, Dryer, & Holekamp, 2011; Cesario & Navarrete, in press; Fessler & Holbrook, 2013; Grinnell, Packer, & Pusey, 1995; McComb, Packer, & Pusey, 1994; Wilson, Britton, & Franks, 2002; Wilson, Hauser, & Wrangham, 2001).

In other words, more than just the primed stimulus information is needed to predict behavior following priming and we should not expect responses to be broadly uniform, at least not when priming socially relevant stimuli. Critics calling for invariant priming effects are effectively arguing that only stimulus information should be important in regulating behavior. This expectation is inconsistent with broader understandings of behavioral regulation and evolutionary understandings of adaptive decision making.

Are Any Priming Effects Real?

So how is it that any one researcher could ever find an effect given what appears to be an overwhelming avalanche of potential variability? It is important to be cautious when answering this question because although there are reasonable answers, there are also unreasonable answers that devolve into excuse making for poor research practices by those of us in the field of priming. My suspicion is that, more often than not, a researcher is accidentally hitting on the right level of an important moderating variable, such as drawing from a population with a low enough mean level of self-monitoring as to obtain a priming effect, or placing participants into cubicles without thinking much of it, or having an experimenter who is low in perceived aggressiveness (allowing participants to respond to the provocation with aggression). And when researchers do not get the "right" combinations of variables, the failures end up in the file drawer. Indeed, this might be what is meant when researchers talk about having "insight" or "intuition" in conducting priming experiments in which they cannot verbalize why they made a decision but knew to make it. In the end, there is really nothing magical about this.

However, there is a danger in this explanation—it may be used to cover up poor research practices and prevent findings that actually are Type I error from ever being corrected. This short circuits the self-correcting nature of science, and it is essential to prevent this from happening. If "researcher degrees of freedom" and questionable research practices (Simmons et al., 2011) are allowing priming researchers to publish findings that are actually Type I error, then it would be a mistake to explain away any subsequent failures and to continue to defend such findings. Along these same lines, priming researchers cannot appeal endlessly to "unknown moderation" without doing the work to provide evidence for such moderation. At some point, the evidence may shift, and it would be more reasonable to conclude that the original effect was wrong or was so specific as to be rendered meaningless. Indeed, the combination of small sample sizes, very large effect sizes, and few direct replications should be a cause for concern. As priming researchers, we must be open to the revision of our beliefs about an effect, not just in theory but in actual practice. The recommendations offered below are designed in part to address these concerns.

Recommendations

At this stage in the development of priming research, it seems apparent to me that the most important step is that we first establish the basic replicability of a given priming effect: Can a single research lab replicate a reported effect multiple times? Only when we have a handle on an effect size of a given effect and we know that the obtained effect is not Type I error can we then move to having other labs attempt replications. At that point, subsequent failures can be understood as holding the promise of an interesting case of moderation. Absent that initial foundation, we will be caught in a perpetual cycle of failures and uncertainties.

Given this, it is the responsibility of priming researchers themselves to provide direct replications of an effect upon initial publication. Having surveyed the literature on behavioral and goal priming, the number of direct replications currently published does not instill a sense of confidence.[5] This is unacceptable for our field and we must do better.

As others have pointed out, replication by multiple independent labs is essential. However it is essential for a given purpose and at the appropriate time. The given purpose is to understand generalizability. The appropriate time is when we have a good enough theory to help us understand the moderating variables that might differ across labs and therefore introduce variability in the original effects. In my assessment, now is not the appropriate time and we cannot yet aim for the purpose of generalizability. This hesitation is due in part to the near-certainty that previous priming effects have been published under methodologies that we are now recognizing as being suboptimal (e.g., questionable research practices, inappropriate analytic techniques: see Judd, Westfall, & Kenny, 2012)—such practices increase Type I error. This point also suggests that priming researchers themselves have to be more conservative and resist making extreme claims about the widespread, highly important, and strong effects of priming. We do not yet know whether this is true, and only widespread direct replications across multiple labs will tell us this information.

It is also important to be clear about what conceptual replications (as opposed to direct replications) can and

cannot tell us, as priming researchers have tended to overstep the limits that can be drawn when discussing conceptual replication. Although a conceptual replication can provide important information, it cannot give the kind of confidence in an effect that direct replication can (nor can it provide information about the effect size of the conceptually replicated effect).

One might be tempted to argue, as priming researchers have (e.g., Bargh, 2012), that a conceptual replication provides additional evidence for the underlying psychological process being tested and therefore instills confidence about other related results. For example, a manuscript may report a study showing that priming "elderly" causes participants to walk more slowly, followed by a study showing that priming "young Black male" causes participants to respond to provocation with greater aggression (Bargh et al., 1996). One might wish to say that the second case of conceptual replication provides good evidence for priming effects in general and decreases the likelihood that the first effect is Type I error. However this is true only if researchers have perfect research and reporting practices. Given the almost certain routine problems of our current practice (such as not being forthright about failed replications), conceptual replications without direct replications will only serve to perpetuate problems in the field. Suppose a manuscript reports three studies, the latter two of which are "conceptual replications" of the first. Perhaps the latter two show interesting moderation of the effect reported in the first study by two different individual difference variables. (This pattern will be familiar to readers of social psychology articles.) Three studies showing an initial effect and subsequently moderation of that effect may appear to be very strong evidence of the researcher's claims. However, rather than being strong evidence, the same set of studies can be generated by including multiple individual difference measures in three separate studies and only reporting such measures when they produce a (false) rejection of the null hypothesis (Simmons et al., 2011). Only direct replication can guard against the use of questionable research practices and other "researcher degrees of freedom," because it would be unlikely to obtain the moderation multiple times if it was simply Type I error. Therefore, it is inappropriate for priming researchers to point to conceptual replications of a given effect as evidence for the robustness, reliability, or generalizability of the priming effect in question.

Conclusion

As priming researchers, we must do better in our methodological practice and provide direct replications when publishing an effect. At the same time, the expectation of widespread invariance in priming effects is inappropriate, and critics of priming must stop drawing strong inferences about replication failures because we are not yet at a stage where we have the quality of theories that render such replications informative. Priming research will get there eventually, but complex social and motivational behaviors take time to understand. Although basic spreading activation effects in cognitive science have been studied for over 50 years, behavioral and goal priming is in its infancy and is, by its nature, more complex.

Theories inform us as to which variables are important and which are unimportant (i.e., which variables can be modified from one research study to the next without consequence). This is one of the reasons why psychology is the hardest science: Social objects are variable in a way that inanimate objects are not (Meehl, 1978). This difficulty is amplified in priming research, given that we do not yet know all the relevant contingencies of the person and context (though of course, many of the problems of this area are shared throughout psychology and beyond; e.g., Gelman, in press).

The linguist Noam Chomsky recently stated:

> Take, say, physics, which restricts itself to extremely simple questions. If a molecule becomes too complex, they hand it over to the chemists. If it becomes too complex for them, they hand it to biologists. And if the system is too complex for them, they hand it to psychologists. (Lawton, 2012)

We should not expect priming effects to be obtained "under all conditions, with all people," a kind of invariance restricted to questions of classical physics. To expect such invariance would be to deny the complexity of human behavior. It is also inconsistent with what we know about the evolved workings of the mind, which center around computing information from different sources (including the physical context) and executing decision rules using this input. Therefore the requirement by some to find a priming paradigm that will work everywhere, every time, and with everyone is a requirement that is simply not reasonable.

There needs to be an appreciation by all parties that every study is merely one data point in the cumulative, ongoing practice of science. For priming researchers, this means to do our part to first establish the replicability of an effect and obtain a reasonable estimate of its effect size and to stop drawing sweeping conclusions about the ubiquity of priming effects before such conclusions are warranted. For critics, this means to give priming research time. The relevant theories are not yet there, but with better practice on the part of priming researchers, we will get there.

Acknowledgments

I thank Daniel Simons, Hal Pashler, Brent Donnellan, and Richard Lucas for productive discussions on this topic. I would especially like to thank Tory Higgins for his early influence on my thinking related to the topics discussed in this article.

Declaration of Conflicting Interests

The author declared no conflicts of interest with respect to the authorship or the publication of this article.

Funding

This material is based upon work supported by National Science Foundation Grant No. 1230281.

Notes

1. *Priming* is of course too broad of a term to be useful. Priming refers to everything from semantic priming in lexical decision tasks (e.g., presenting *doctor* makes it faster for participants to identify *nurse* as a word; Meyer & Schvaneveldt, 1971; Neely, 1977) to goal priming (e.g., priming *achievement* makes call center employees solicit more money from donors; Shantz & Latham, 2009). To be more specific, it is the research showing priming effects on social behavior and goal-directed behavior that has been targeted most for criticism. Such research has gone under the names "behavioral priming," "goal priming," "automatic social behavior," and "social priming." For simplicity, I use the term "priming" throughout, but the arguments in this article apply exclusively to priming effects on social and goal-related behavior.

2. Of course, even with this simple example there are complications. For example, participants in American University subject pools at the beginning of a semester are more conscientious than those at the end of a semester (Witt, Donnellan, & Orlando, 2011). If the correlation between the two measures in this hypothetical example is moderated by conscientiousness, then one might only obtain the correlation when sampling from the first, and not the second, half of the semester. Thus doing the same thing might even be dependent on sampling during the same month of the year (or even time of day; Bodenhausen, 1990), a fact that does not diminish the importance of or interest in the effect.

3. Several reviewers of this article emphasized that theory development has occurred and that the pace of such development is appropriate, with several comprehensive models emerging within the short couple decades of priming research: Jonas and Sassenberg's (2006) behavioral response account, Cesario et al.'s (2006) motivated preparation to interact account, Wheeler, DeMarree, and Petty's (2007) active self account, Loersch and Payne's (2011) situated inference model, and Schröder and Thagard's (2013) unified computational model. I thank these reviewers for this reminder. I hesitate, however, to agree with the suggestion by these reviewers that I should be more optimistic about these advances and the state of the field. My hesitation is informed, first, by the variable quality of the methodologies and statistical analyses of past priming research. Second, and more important, is the continued resistance by some priming researchers to improve the quality of their research with even modest changes to their methodological practice. For example, as of this writing, there is an in-press manuscript from a leading priming researcher reporting a single study with $n = 15$ per cell and $p = .055$ effect.

4. This discussion is reminiscent of the debate surrounding the effects of endostatins on tumor growth (O'Reilly et al., 1997; Rowe, 1999), a question that experienced years of replication failures and successes and ultimately uncovered that replication outcomes hinged on very nuanced but unanticipated features of the experimental protocol ("Folkman says it took him years to perfect his techniques, which rely on factors such as the amount of material in the syringe, the gauge of the needle, where you inject the mice, and the temperature of the room that houses the animals" Cohen, 1999, p. 1251).

5. It is interesting to note that the paper generating more criticism than any other priming paper contains one of the few direct replications in the field (Studies 2a and 2b of Bargh et al., 1996).

References

Asendorpf, J. B., Conner, M., de Fruyt, F., de Houwer, J., Denissen, J. J. A., Fiedler, K., . . . Wicherts, J. M. (2013). Recommendations for increasing replicability in psychology. *European Journal of Personality*, *27*, 108–119. doi:10.1002/per.1919

Bargh, J. A. (2012, March). Angry birds. *Psychology Today*. Retrieved from https://docs.google.com/document/d/1wuu8URArgZusJELXF5j4xpM26ESkFfOveYoGKBf0CHo/edit?pli=1

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*, 230–244. doi:10.1037/0022-3514.71.2.230

Benson-Amram, S., Heinen, V. K., Dryer, S. L., & Holekamp, K. E. (2011). Numerical assessment and individual call discrimination by wild spotted hyenas, *Crocuta crocuta*. *Animal Behaviour*, *82*, 743–752.

Blair, I. V., Judd, C. M., & Fallman, J. L. (2004). The automaticity of race and Afrocentric facial features in social judgments. *Journal of Personality and Social Psychology*, *87*, 763–778.

Blanchard, D. C., Hynd, A. L., Minke, K. A., Minemoto, T., & Blanchard, R. J. (2001). Human defensive behaviors to threat scenarios show parallels to fear- and anxiety-related defense patterns of non-human mammals. *Neuroscience & Biobehavioral Reviews*, *25*, 761–770.

Blanchard, R. J., Flannelly, K. J., & Blanchard, D. C. (1986). Defensive behaviors of laboratory and wild *Rattus norvegicus*. *Journal of Comparative Psychology*, *100*, 101–107.

Bodenhausen, G. V. (1990). Stereotypes as judgmental heuristics: Evidence of circadian variations in discrimination. *Psychological Science*, *1*, 319–322. doi:10.1111/j.1467-9280.1990.tb00226.x

Bos, P. A., Panksepp, J., Bluthé, R. M., & van Honk, J. (2012). Acute effects of steroid hormones and neuropeptides on human social-emotional behavior: A review of single administration studies. *Frontiers in Neuroendocrinology*, *33*, 17–35. doi:10.1016/j.yfrne.2011.01.002

Brown, C. M., & McConnell, A. R. (2009). When chronic isn't chronic: The moderating role of active self-aspects. *Personality and Social Psychology Bulletin*, *35*, 3–15. doi:10.1177/0146167208324675

Bry, C., Follenfant, A., & Meyer, T. (2008). Blonde like me: When self-construals moderate stereotype priming effects on intellectual performance. *Journal of Experimental Social Psychology*, *44*, 751–757. doi:10.1016/j.jesp.2007.06.005

Cesario, J., & Jonas, K. J. (2013). *The role of in-group coalitions in action preparation: A resource computation framework for automatic responses*. Manuscript in preparation.

Cesario, J., & McDonald, M. M. (2013). Bodies in context: Power poses as a computation of action possibility. *Social Cognition*, *31*, 260–274.

Cesario, J., & Navarrete, C. D. (in press). Perceptual bias in threat distance: The critical roles of ingroup support and target evaluations in defensive threat regulation. *Social Psychological & Personality Science*.

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*, 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*, 893–910. doi:10.1037/0022-3514.90.6.893

Cohen, J. (1999). Behind the headlines of endostatin's ups and downs. *Science*, *283*, 1250–1251.

Custers, R., & Aarts, H. (2005). Positive affect as implicit motivator: On the nonconscious operation of behavioral goals. *Journal of Personality and Social Psychology*, *89*, 129–142. doi:10.1037/0022-3514.89.2.129

DeMarree, K. G., Wheeler, S. C., & Petty, R. E. (2005). Priming a new identity: Self-monitoring moderates the effects of nonself primes on self-judgments and behaviors. *Journal of Personality and Social Psychology*, *89*, 657–671. doi:10.1037/0022-3514.89.5.657

Dijksterhuis, A., Aarts, H., Bargh, J. A., & van Knippenberg, A. (2000). On the relation between associative strength and automatic behavior. *Journal of Experimental Social Psychology*, *36*, 531–544. doi:10.1006/jesp.2000.1427

Fessler, D. M. T., & Holbrook, C. (2013). Friends shrink foes: The presence of comrades decreases the envisioned physical formidability of an opponent. *Psychological Science*, *24*, 797–802. doi:10.1177/0956797612461508

Gelman, A. (in press). The connection between varying treatment effects and the crisis of unreplicable research: A Bayesian perspective. *Journal of Management*.

Grinnell, J., Packer, C., & Pusey, A. E. (1995). Cooperation in male lions: Kinship, reciprocity or mutualism? *Animal Behaviour*, *49*, 95–105.

Hagiwara, N., Kashy, D. A., & Cesario, J. (2012). The independent effects of skin tone and facial features on Whites' affective reactions to Blacks. *Journal of Experimental Social Psychology*, *48*, 892–898. doi:10.1016/j.jesp.2012.02.001

Heidbreder, E. (1933). *Seven psychologies*. New York, NY: Appleton-Century-Crofts.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral & Brain Sciences*, *33*, 61–135. doi:10.1017/S0140525X0999152X

Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 133–168). New York, NY: Guilford Press.

Ioannidis, J. (2005). Why most published research findings are false. *PLoS Medicine*, *2*(8), e124. Retrieved from http://www.plosmedicine.org/article/info%3Adoi%2F10.1371%2Fjournal.pmed.0020124

Jonas, K. J., & Sassenberg, K. (2006). Knowing how to react: Automatic response priming from social categories. *Journal of Personality and Social Psychology*, *90*, 709–721.

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, *103*, 54–69. doi:10.1037/a0028347

Katzko, M. W. (2006). A study of the logic of empirical arguments in psychological research: "The automaticity of social behavior" as a case study. *Review of General Psychology*, *10*, 210–228. doi:10.1037/1089-2680.10.3.210

Lawton, G. (2012, March). Everything was a problem and we did not understand a thing: An interview with Noam Chomsky. *New Scientist, 2856*. Retrieved from http://www.newscientist.com/article/mg21328560.200-noam-chomsky-meet-the-universal-man.html

Loersch, C., Aarts, H., Payne, B. K., & Jefferis, V. E. (2008). The influence of social groups on goal contagion. *Journal of Experimental Social Psychology*, *44*, 1555–1558. doi:10.1016/j.jesp.2008.07.009

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*, *6*, 234–252.

Ma, D. S., & Correll, J. (2011). Target prototypicality moderates racial bias in the decision to shoot. *Journal of Experimental Social Psychology*, *47*, 391–396.

McComb, K., Packer, C., & Pusey, A. (1994). Roaring and numerical assessment in contests between groups of female lions, *Panthera leo*. *Animal Behaviour*, *47*, 379–387.

Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology*, *46*, 806–834.

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*, 227–234.

Mussweiler, T., & Förster, J. (2000). The sex→aggression link: A perception-behavior dissociation. *Journal of Personality and Social Psychology*, *79*, 507–520. doi:10.1037//0022-3514.79.4.507

Mussweiler, T., & Neumann, R. (2000). Sources of mental contamination: Comparing the effects of self-generated versus externally provided primes. *Journal of Experimental Social Psychology*, *36*, 194–206. doi:10.1006/jesp.1999.1415

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, *106*, 226–254.

O'Reilly, M. S., Boehm, T., Shing, Y., Fukai, N., Vasios, G., Lane, W. S., . . . Folkman, J. (1997). Endostatin: An endogenous inhibitor of angiogenesis and tumor growth. *Cell*, *88*, 277–285.

Pashler, H., & Wagenmakers, E. J. (2012). Editors' introduction to the special section on replicability in psychological science: A crisis of confidence? *Perspectives on Psychological Science*, *7*, 528–530. doi:10.1177/1745691612465253

Rowe, P. M. (1999). What is all the hullabaloo about endostatin? *The Lancet*, *353*, 732.

Schröder, T., & Thagard, P. (2013). The affective meanings of automatic social behaviors: Three mechanisms that explain priming. *Psychological Review*, *120*, 255–280. doi:10.1037/a0030972

Sell, A., Tooby, J., & Cosmides, L. (2009). Formidability and the logic of human anger. *Proceedings of the National Academy of Sciences, USA*, *106*, 15073–15078.

Shantz, A., & Latham, G. P. (2009). An exploratory field experiment of the effect of subconscious and conscious goals on employee performance. *Organizational Behavior and Human Decision Processes*, *109*, 9–17. doi:10.1016/j.obhdp.2009.01.001

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*, 1359–1366. doi:10.1177/0956797611417632

Spellman, B. A. (2013). Introduction to the special section on advancing science. *Perspectives on Psychological Science*, *8*, 412–413.

Tooby, J., & Cosmides, L. (2005). Conceptual foundations of evolutionary psychology. In D. M. Buss (Ed.), *Handbook of evolutionary psychology* (pp. 5–67). Hoboken, NJ: Wiley.

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*, *11*, 234–261.

Wilson, M. L., Britton, N. F., & Franks, N. R. (2002). Chimpanzees and the mathematics of battle. *Proceedings of the Royal Society B: Biological Sciences*, *269*, 1107–1112.

Wilson, M. L., Hauser, M. D., & Wrangham, R. W. (2001). Does participation in intergroup conflict depend on numerical assessment, range location, or rank for wild chimpanzees? *Animal Behaviour*, *61*, 1203–1216.

Witt, E. A., Donnellan, M. B., & Orlando, M. J. (2011). Timing and selection effects within a psychology subject pool: Personality and sex matter. *Personality and Individual Differences*, *50*, 355–359. doi:10.1016/j.paid.2010.10.019