Can the goal of honesty be primed?

Harold Pashler [a,*], Doug Rohrer [b], Christine R. Harris [a]

[a] University of California, San Diego, USA
[b] University of South Florida, USA

HIGHLIGHTS

• We attempted to replicate the "Honesty Priming" effect of Rasinski et al. (2005).
• Two direct replications failed to reproduce the effect.
• An additional conceptual replication using a broader outcome measure also failed.

ABSTRACT

In a simple study involving 64 participants, Rasinski, Visser, Zagatsky, and Rickett (2005) reported that requiring people to make semantic judgments involving four words related to honesty (embedded among other words) increased the likelihood that they would later admit to having engaged in problematic alcohol-related behaviors (e.g., drinking to the point of blackout). If valid, this honesty-priming effect would offer a powerful intervention to improve the validity of self-report data in many different contexts. To determine whether the effect is repeatable, we first attempted two replications using the same materials, tasks, and measures used by Rasinski et al. Experiment 1 repeated the study with a sample of 150 students. No priming effects were observed here, nor in a follow-up study using adults recruited on the web (Experiment 2). Experiment 3 used the same priming manipulation together with a more refined measure of response candor (derived from Paulhus, 1991). Again, the honesty-related primes had no detectable effects.

© 2013 Elsevier Inc. All rights reserved.

Introduction

The term 'priming' encompasses a broad range of different behavioral effects that can be elicited by giving people a relatively incidental or minimal exposure to words, pictures, or other stimuli. Some priming effects involve cognitive and perceptual changes, typically biases toward perceiving stimuli as being related to the priming materials (e.g., Johnston & Hale, 1984; Meyer & Schvaneveldt, 1971; Schvaneveldt & McDonald, 1981). In recent years, however, a far more diverse set of priming effects have been reported in the psychological literature. These effects involve changes in how well people perform complex tasks, higher-level judgments about many kinds of topics, and even the choice of actions or styles of actions. While the literature contains quite a few conceptual replications of these broader types of priming effects, some recent attempts to replicate the results directly have reported failures (Doyen, Klein, Pichon, & Cleeremans, 2012; Pashler, Coburn, & Harris, 2012; Shanks et al., 2013).

Honesty priming

The focus of the current article is a goal-priming finding that would appear to have great potential for practical application, as well as considerable scientific interest. This effect was reported in *Journal of Experimental Social Psychology* by Rasinski et al. (2005). Rasinski et al. (2005) reported that by merely exposing people to some honesty-related words in a synonym judgment task, they were able to elicit a higher than normal level of candor from subjects in a subsequent questionnaire asking about embarrassing behaviors (alcohol abuse). Subjects saw six target words; for subjects in the honesty condition, four of these were related to the goal of being honest (e.g., *honest*, *genuine*). Each of these target words were presented with three roughly synonymous terms (e.g., *open*, *sincere*, and *truthful*), and the subject's task was to select which one of these three words seemed most similar to each target word. Participants in a control condition performed the same task with six target words unrelated to honesty (e.g., *blend*, *common*). After subjects completed the priming task, they completed an ostensibly unrelated questionnaire, which asked them some questions about past drinking behaviors (binge drinking, blackouts, etc.). The proportion of subjects who admitted to problematic levels of drinking was higher for subjects who had been primed with the honesty-related terms.

* Corresponding author.
  E-mail address: hpashler@gmail.com (H. Pashler).

If this effect is robust, it would seem to have considerable potential impact in practical areas. Collection of self-report data from people is a commonplace and costly activity in domains ranging from marketing research to public health and opinion polling. A very low-cost method of increasing candor from respondents would have major significance.

Given the potential importance of this effect, and the emergence of considerable skepticism about goal priming findings (Yong, 2012a, 2012b), the Rasinski et al. (2005) study appeared to be a worthwhile result to try to verify. The present authors have had difficulty replicating a number of goal priming results, but we suspected in advance that this particular effect was more likely to be confirmable because it involved a priming manipulation far less minimal and incidental than many priming manipulations in the literature. After all, subjects did not merely read words like "honesty"—they contemplated different aspects of personal integrity.

The original study had some features that might raise statistical concerns, however. In addition to the drinking questions, Rasinski et al. (2005) also had subjects answer questions about academic cheating. For half of the subjects, the drinking questions preceded the cheating questions. For these subjects, exposure to the honesty primes caused a significant increase in subjects' likelihood of problematic drinking ($t(60) = 4.67$, $p < .001$, $d = 1.21$). However, the other half of the subjects answered the cheating questions *before* the alcohol questions, and for them, the honesty primes did not affect their responses to the alcohol questions ($t(60) = 0.96$). Honesty primes did not significantly affect responses to the cheating questions (although there was a very weak trend, $p = .19$). Reasonably enough, Rasinski et al. concluded that the lack of a priming effect on the alcohol questions when these questions *followed* the cheating questions might be due to the fact that priming effects weaken rapidly. However, they did not offer any explanation for why honesty priming did not affect the questions about cheating, even though one would suppose that students would be ashamed of cheating as well as drinking.

Given that Rasinski et al. (2005) found a priming effect in only 1 of the 4 conditions (alcohol questions that appear immediately after the prime), this is the only effect we tried to replicate. Experiment 2 repeated the study with a diverse sample of approximately 150 adults tested online. Experiment 3 examined yet another online sample of about 150 people using the same priming manipulation used by Rasinski et al. but different questions, designed to provide a more refined measure of candor. (A final study very similar to Experiment 3 is reported in the Online Supplement.)

Rasinski et al. (2005) noted that their results could potentially have reflected demand effects, but argued that this was unlikely. In order to assess this possibility, subjects in our studies were asked several additional questions at the end of the study to determine if they might have guessed the hypothesis. This issue will be discussed further in the General discussion.

Experiment 1

Method

Participants and design

To provide more statistical power than in the original study, we decided in advance to test approximately 2.5 times the number of participants used by Rasinski et al. (2005). In Experiment 1, the participants were college students, as in Rasinski et al. A total of 150 undergraduate subjects from the UCSD Psychology Department Subject Pool completed this experiment, participating in partial fulfillment of a course requirement.

Materials

The materials used by Rasinski et al. (2005) were obtained from the original authors.[1] This included the synonym judgment task

materials for both honesty and neutral conditions (see Appendix A), the alcohol questionnaire, and instructions provided to subjects for both the synonym judgment task and the drinking questionnaire.

The primes and questions were printed on paper. The instructions for this task provided a vague paragraph-long psycholinguistically-oriented context ("Communication is a complicated process and even simple words can have slightly different meanings to people. We're interested in how people think about particular words, or what those words <u>mean</u> to people.") The alcohol questionnaire was labeled a "Health Behaviors Questionnaire" and printed in a different font type than the judgment task. The instructions for the drinking questions stated, "The following questions are confidential. Your responses will be kept in strictest confidence and will not be stored in a file containing your name or other contact information." The questions inquired about binge drinking, missing class due to drinking the previous evening, drinking enough to have difficulty remembering activities, and driving after consuming two or more alcoholic beverages. A final questionnaire was included at the very end, but on a half sheet of paper. All 3 forms were stapled together, folded into thirds, shuffled, and then a number was placed on the paper (which was later used to pair with the subject ID). Each questionnaire was placed in an envelope to prevent experimenters from being able to see the forms. The outside of the envelope contained a random letter (i.e. A–Z). The envelopes were shuffled again and placed in a box, so that subjects could reach in and grab one form. To ensure that experimenters would not be exposed to the materials during the study, the person who prepared the forms was not the same person who ran the participants.[2]

Procedures

The procedures generally followed those used by Rasinski et al. (2005) except where noted. As in Rasinski et al., the study was presented as consisting of two separate elements; participants were told that the study involved "two questionnaires" with a total of 10 questions and that it would take no more than 5 min to complete both.

For the synonym judgment task, the subject read instructions that said that they should "read each word carefully and indicate which of the three words seems most similar to the first word." In the neutral condition, the target words were *peaceful, blend, common, prepared, plain* and *avid*. In the honesty condition, the words were *honest, secure, common, genuine, plain* and *correct*. For each word, the subject circled one of three roughly synonymous alternatives (e.g., for *genuine* the alternatives offered were *real, straightforward*, and *true*.)

When subjects arrived at the lab, they were asked to turn off any cellphone they might have and were then taken into individual sound-attenuating booths to begin reading over the consent form. The experimenter had each subject select one envelope from the box provided and wrote down the letter on the outside of the envelope. Subjects were instructed not to begin until the experimenter had left the room and reminded not to write their name on any of the questionnaires. The subjects were also told that the experimenter was not permitted to look at the forms, and therefore, subjects should write a note on their feedback sheet if they had any questions or comments for the researcher. This further ensured that experimenter had neither exposure to the material nor awareness of the hypotheses. (No similar precautions were described by Rasinski et al. (2005).)

After completion of the survey, subjects were instructed to place their forms back into the envelope and to place the envelope into a locked survey box that was provided for them. They were then provided with a printed debrief sheet to read over and told that they could contact the researcher if they needed more detailed information.

---

[1] We are grateful to Kenneth Rasinski for providing these materials.

[2] An exception being 3 subjects who were run by the lab manager when the regular experimenters were not available due to illness (the lab manager had no way of knowing which condition the subject was in, however).

Results and discussion

Data were obtained from 150 subjects. One of the these participants said, in response to the request to speculate about the purpose of the study, "Maybe the word-matching task in the beginning was meant to prime me to be honest, sincere, straightforward, etc. in the next set of personal questions." This subject's data were excluded from further analysis. The other 149 subjects' responses did not reveal any awareness of the hypothesis.

The mean proportion of drinking questions to which the subject responded "yes" was 0.355 (SD = .368) for the Honesty condition and 0.290 (SD = .334) for the Neutral condition. The difference was not significant, $t(147) = 1.12$, $p = .26$, $d = 0.184$. The results provided no evidence for any reliable effect of priming honesty on subjects' responses to questions about alcohol use.

Experiment 2

To further examine the validity of the honesty priming effect reported by Rasinski et al., the study was repeated online using a diverse panel of adult subjects. While online testing differs from the original procedure, it has the advantage that it eliminates any potential for unintended social influence on the participant (already a rather remote scenario given the precautionary procedures described above).

Method

Participants and design

Subjects were drawn from the Internet Subject Panel of the UCSD Learning, Attention, and Perception Lab, composed of a diverse sample of adults who have previously participated in cognitive studies conducted by members of this lab. Potential subjects were invited by email, and subjects who agreed to participate were randomly assigned to either the honesty condition or the neutral condition. Subjects in this pool are generally paid about $10.00 per hour, with a slightly greater rate of pay for short studies. The email offered subjects an opportunity to "participate in two brief studies" taking about 5 min, in return for which they would receive $1.50. (Participants in this pool receive invitations with vague descriptions of experiments so that they cannot self-select based on their interests or other personal factors that might create an unrepresentative sample of the panel.) A total of 152 subjects (65% female; average age 38 years) completed the study.

Materials

The materials from Experiment 1 were presented on computer screen. Different fonts were used to display materials for the synonym judgment task and the drinking questionnaire, to encourage subjects to view these as two separate studies.

Procedure

The online program was designed to randomly assign subjects to conditions using a PHP script designed to produce roughly equal numbers of subjects in each of the two conditions. (Due to a programming error, the first 38 subjects were all assigned to the Neutral condition. This error was corrected as soon as it was detected, and the remaining subjects were randomly assigned so that the entire sample would end up with approximately equal numbers in each group — this did not affect the results as described below.) Immediately after the prime condition, subjects answered the questions on alcohol use. After completion of the alcohol questionnaire, subjects were asked to "try to guess the purpose of these studies" to which they were allowed to type in a response.

Results and discussion

Data were obtained from 70 participants in the honesty condition and 82 participants in the neutral condition. The mean proportion of alcohol questions to which the subject responded "yes" were 0.407 (SD = .381) for the honesty condition and 0.457 (SD = .359) for the neutral condition (see Fig. 1, which also shows the results from Rasinski et al., 2005). The effect of prime type was not significant, $t(150) = -.83$, $p = .406$, $d = -0.136$.

None of the 152 subjects showed any evidence of having inferred the purpose of the study. Specifically, none noted the presence of honesty-related words in the semantic task, nor did any suggest that the purpose of the study might have been to see whether content in the semantic task affected their honesty and candor in the questionnaire. (Finally, to ensure that the program randomization error did not affect our findings, we conducted an additional analysis of our data excluding the initial 38 subjects. There was still no hint of an effect of prime condition: $t(112) = -0.27$, $p = .789$. The observed means for this subset of 114 subjects were 0.407 (SD = .381) for the Honesty Condition vs. 0.426 (SD = .360) for the Neutral Condition.) The results do not find support for the existence of any effect of priming honesty on responses to questions about alcohol abuse in this adult sample.

Discussion of Experiments 1 and 2

The direct replication attempt in Experiment 1 and the similar study with a more diverse sample in Experiment 2 do not replicate the priming effect reported by Rasinski et al. (2005). A reviewer of an earlier version of the present article noted that the average scores on the drinking behaviors for the unprimed subjects in our studies (29%) were somewhat higher than the score for the unprimed subjects in Rasinski et al. for the drinking-questions-first condition (18%). Thus, the reviewer suggested that perhaps our subjects were less inhibited about reporting drinking to start with (i.e., many were *already* being completely honest even without the primes, and thus priming did not cause them to report a higher level of drinking). This account seemed somewhat implausible to us for a number of reasons.

First, half of the participants in Rasinski and colleagues' study were given the drinking questions after they answered the cheating questions; for their subjects as a whole, the mean response rate was 26%, which is nearly the same as our college sample (29%). Moreover, relative to our sample, the *primed* subjects in Rasinski et al. (2005) reported substantially higher rates on the drinking questionnaire (67% and 43% for the drinking questions first and last groups, respectively).

Second, given the demographic heterogeneity of University of California students, it seems to us implausible a priori that there would be any stark differences between how people in the two samples would feel about revealing negative drinking behaviors. (In their Introduction, Rasinski et al. (2005) describe studies showing that people of different ages in varied societies tend to underreport very heavy drinking, and it seems doubtful that the University of California would represent an exception to this.)

Third, while the *average* level of drinking reported by our participants differed from that of the participants in Rasinski et al. (2005), there is a high degree of overlap between the distributions of participants' responses in these various studies. Even when variation in one variable strongly moderates the effect of another variable, the sort of natural variation seen within different samples is usually gentle, precluding dramatic interactions from occurring.[3] Thus, the idea that priming would have a strong effect in one sample and none in another group simply due to differences in the natural distribution of a putative moderating variable seems to us rather implausible.

---

[3] As McClelland and Judd (1993) pointed out, there was considerable frustration in industrial and organizational psychology in the 1970s and 1980s as interaction and moderation effects that had been found robust in experimental studies (in which extremes of the moderator were randomly assigned) failed to reach statistical significance in field studies (in which the presumed moderator had a broad distribution). These authors analyzed this phenomenon in detail using simulations.
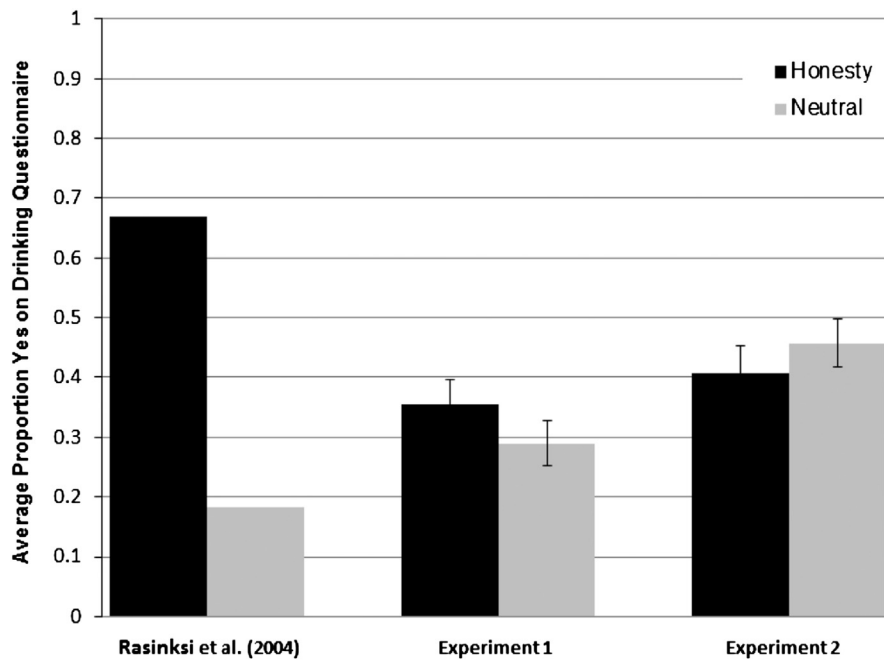
Fig. 1. Mean responses to the drinking questions for the study reported by Rasinski et al. (2005; Drinking Questions First Group) and in Experiments 1 and 2. Error bars show standard error of the mean (not available for Rasinski et al.).

Experiment 3

Whatever may be causing the differences between the results reported above and those reported by Rasinski et al. (2005), it is a psychometrically unfortunate feature of the design of all three studies mentioned to have attemped to measure respondents' candor with just a few questions all confined to a single domain (drinking). Fortunately, personality and psychopathology researchers have long been interested in methods for assessing people's honesty in completing questionnaires, and have developed refined measures of candor (or its absence). Therefore, as an additional (and in our judgment, improved) study, we looked at how the honesty priming manipulation would affect scores on Paulhus (1991) Balanced Inventory of Desirable Responding (BIDR). The BIDR contains two subscales. The first, the Impression Management (IM) subscale, is seen as representing conscious deception in order to present a positive impression. The second, the Self-Deceptive Enhancement (SDE) scale, is viewed as representing largely unconscious self-deception in which an individual denies "foibles" (e.g., saying, "I never regret my decisions.").

Method

Participants and design

A sample of 151 adult participants was recruited from Amazon Mechanical Turk. The sample was restricted to participants from the United States who had approval ratings of 80% or greater on previous Mechanical Turk jobs.

Procedure

The procedures followed Experiment 2 closely, except that instead of the drinking questions, subjects were given the BIDR. The BIDR's response scale with its verbal labels was displayed on top of the page, and the subject used radio buttons to choose a response for each item.

Results

None of the subjects articulated any suspicion that the synonym judgment task was intended to alter performance on the BIDR questionnaire. For the SDE subscale, there was no significant difference between the honesty condition (M = 5.10, SD = 4.12) and the neutral condition (M = 5.69, SD = 4.07), t(149) = −.882, p = .380. For the IM Subscale, there was also no significant difference between the honesty condition (M = 5.39, SD = 4.04) and the neutral condition (M = 5.92, SD = 4.29), t(149) = −.779, p = .437. Results are shown in Fig. 2. The Cronbach's Alphas for the questionnaire subscales were .817 and .824 (which compare well with norming data reported by Paulhus, 1991, suggesting that these Mechanical Turk participants were probably no less careful in their responding than were people in the norming sample used by Paulhus).

Discussion of Experiment 3

This study used a refined inventory of honest versus deceptive reporting, and the same honesty priming manipulation used by Rasinski et al. (2005). Again, the priming intervention had no significant effect (or even a strong trend) either way on the two subscales. (We also made an additional attempt to replicate the result using an English-speaking sample of Mechanical Turk users from India, and here too, the results showed no effects; see Online Supplement.)

General discussion

The present work described a sustained attempt to validate the honesty priming finding reported by Rasinski et al. (2005). All three experiments reported here involved samples of approximately 150 people per study and pre-planned analysis methods. None of the studies provided any evidence that honesty priming increases candor of responding. A fourth study (reported in the online supplement) also did not support the findings of Rasinski et al. (2005); these four studies represent the entirety of the data collection that we have undertaken relating to honesty priming.

Considering just the results of the two replication attempts that used the same manipulation and the same dependent variable as the original work (i.e., Experiments 1 and 2), the present sample involves 301 participants. For this combined sample, the mean drinking behavior scores were 0.380 and 0.377 for honesty and neutral conditions, respectively. There was no significant difference, t(299) = .067,
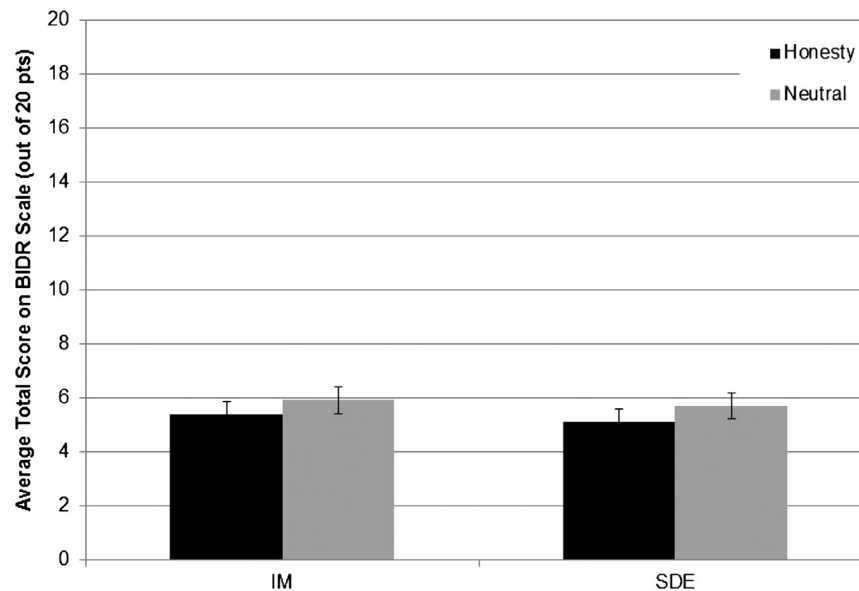
Fig. 2. Mean scores on Impression Management (IM) and Self-Deceptive Enhancement (SDE) subscales of Balanced Inventory of Desirable Responding (BIDR) as a function of the priming manipulation condition (Experiment 3).

$p = .947$, and the overall effect size on the combined sample was $d = 0.008$ (with a 95% confidence interval ranging from $-.219$ to $+.234$). The results argue against the existence of a large (or even medium-sized) effect of honesty primes and provide no support for the idea that honesty priming can enhance the validity of self-report data.

Some researchers have argued that the analysis of a direct replication attempt should include a statistical test of the difference between the effect size observed in an original study and the effect size found in a replication attempt (e.g., Srivastava, 2012). In this case, we were unable to perform this test because Rasinski et al. (2005) did not report measures of variability (such as standard deviations, standard errors, or error bars). However, given the very large effect size reported in their original study ($d = 1.21$ for the alcohol questions first group) and the confidence interval for the effect size found in our combined sample (-0.219 to $+0.234$), we suspect the difference between the priming effect reported in the original study and the null effect found in the combination of Experiments 1 and 2 is almost certainly a statistically reliable one.

Possible explanations for the discrepancy

Based on a reading of Rasinski et al. (2005), we began this work with the thought that a failure to replicate, should it occur, might reflect possible demand effects in the original study. In the study reported by Rasinski et al. no debriefing questions were asked, leading Rasinski et al. to concede that their findings could possibly have resulted from subjects guessing the intent of the experiment and then responding accordingly. However, they noted that in a separate experiment (which they did not report) debriefing questions were included, and that "with literally one exception, none of the roughly 130 participants have said anything that indicated any awareness of the experimental hypothesis or any suspicion about the true purpose of the word meaning task" (p. 6). (They did not report on the drinking behavior results of this study.) The present studies included debriefing questions as described above, and the results (with only one of the participants in Experiments 1–3 guessing the hypothesis) seem fully consistent with the comments by Rasinski et al. Thus, a demand effect

does not seem like a likely cause for the positive finding of Rasinski and colleagues.

So what general conclusions can be drawn? Of course, it is always possible that some unintended and unnoticed difference between the original study and the replication attempts reported here was responsible for the different outcomes. For one thing, the participants were not identical: the original study used University of Chicago students, whereas our Experiment 1 used University of California students. The materials were as similar as possible to the originals as could be arranged, even using the same text provided by the original investigators. However, the original study stated that "participants sat alone at a table in the student center and completed the survey" whereas subjects in Experiment 1 were run alone in booths in a laboratory. Moreover, the interaction between experimenter and subject was explicitly designed to ensure that experimenters would remain blind; no similar precautions were described by Rasinski et al. (2005). It seems to us that all such very tiny differences are highly unlikely to have been crucial (and of course if they are, honesty priming is not likely to offer any practical way to enhance self-report data).

A different possibility (which seems more plausible to the current authors) is that the Rasinski et al. (2005) result may simply reflect a Type 1 error.[4] Similar possibilities exist for other goal priming results whose validity has recently come into question (e.g., Doyen et al., 2012; Pashler et al., 2012; Shanks et al., 2013). The likely rate of such errors in the goal priming literature—and, more broadly, scientific literatures generally—is a point of much current controversy and speculation (Bargh, 2012; Bower, 2012; Ioannidis, 2005). Due to the "file drawer problem" (the notorious tendency for scientific journals to selectively publish positive outcomes; cf. Ioannidis, 2005; Rosenthal, 1979) a published literature can easily provide a very misleading picture of reality. For example, if goal priming results (many

---

[4] The absence of an effect of honesty primes on the cheating reports in their study encourages this line of thought.

of which have received extensive press attention; see Bower, 2012) have attracted a great many investigators due to their counterintuitive nature and simplicity, then an enormous number of attempts to obtain such effects could have been made over the past several decades. At least 5% of these would be expected to have "succeeded" just by chance. Moreover, if some authors adopt practices that exploit hidden degrees of freedom (termed "p-hacking" by Simmons, Nelson, & Simonsohn, 2011), then the proportion of studies yielding false-positive results could have been substantially higher.

Thus, the fact that there are many dozens (and perhaps more than 100) goal priming results in the published literature does not necessarily imply that these effects are real and reproducible (Bakker, van Dijk, & Wicherts, 2012; Pashler & Harris, 2012). This frustrating indeterminacy—which is by no means confined to goal priming—reflects both the rarity of direct replication attempts and the disinclination of journals and investigators to publish null results (factors that are increasingly recognized as injurious to the credibility of almost every area of science; Ioannidis, 2005; Pautasso, 2010; Young, Ioannidis, & Al-Ubaydli, 2008). At this point, then, it is not clear whether goal priming is a counterintuitive but real phenomenon (fragile in some fashion that is not presently understood)—or, instead, may provide a dramatic illustration of how publication bias can impede scientific understanding (particularly when "conceptual replications" are taken to be an adequate substitute for direct replications; cf. Pashler & Harris, 2012). This question can only be cleared up by additional efforts aimed at directly replicating additional published goal priming results with good statistical power and clearly described protocols.[5] Fortunately, interest in carrying out direct replications in the goal priming area appears to be growing rapidly now (Bower, 2012; Yong, 2012b), and thus it seems realistic to hope that some clarity on this point may emerge over the coming years.

Acknowledgments

Appendix A

Words used for the synonym judgment task which constituted the priming manipulation in this study. Materials were first used by Rasinski et al. (2005). For each condition, there are six trials, utilizing 24 words. Participants were shown the first word in each set of four words, and asked to select which of the remaining three words "seems most similar to the first word". For example, in the Honesty Priming Condition, the subject would indicate which of the set *open*, *sincere* or *truthful* seemed more similar to *honest*.

Honesty priming condition

Honest	Open	Sincere	Truthful
Secure	Safe	Comfortable	Protected
Common	Frequent	Routine	Average
Genuine	Real	Straightforward	True
Plain	Neutral	Simple	Basic
Correct	Actual	Straight	Accurate

Neutral condition

Peaceful	Calm	Quiet	Passive
Blend	Mix	Combine	Merge
Common	Frequent	Routine	Average
Prepared	Equipped	Ready	Set
Plain	Neutral	Simple	Basic
Avid	Eager	Ready	Enthusiastic

Appendix B. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jesp.2013.05.011.

References

Bakker, M., van Dijk, A., & Wicherts, J. M. (2012). The rules of the game called psychological science. *Perspectives on Psychological Science, 7*, 543–554.

Bargh, J. A. (2012). *Priming effects replicate just fine, thanks.* Psychology Today (blog http://www.psychologytoday.com/blog/the-natural-unconscious/201205/priming-effects-replicate-just-fine-thanks, published May 11, 2012).

Bower, B. (2012). The hot and the cold of priming. *Science News, 181*.

Doyen, S., Klein, O., Pichon, C. -L., & Cleeremans, A. (2012). Behavioral priming: It's all in the mind, but whose mind? *PLoS One, 7*(1), e29081. http://dx.doi.org/10.1371/journal.pone.0029081.

Ferguson, C. J., & Heene, M. (2012). A vast graveyard of undead theories: Publication bias and psychological science's aversion to the null. *Perspectives on Psychological Science, 7*, 555–561.

Ioannidis, J. P. A. (2005). Why most published findings are false. *PLoS Medicine, 2*, e124.

Johnston, J. C., & Hale, B. L. (1984). The influence of prior context on word identification: Bias and sensitivity effects. *Attention & Performance X, 243–255*.

McClelland, G. H., & Judd, C. M. (1993). Statistical difficulties of detecting interactions and moderator effects. *Psychological Bulletin, 114*, 376–390.

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.

Pashler, H., Coburn, N., & Harris, C. R. (2012). Priming of social distance? Failure to replicate effects on social and food judgments. *PLoS One, 7*(8) (doi: http://dx.doi.org/10.1371/journal.pone.0042510).

Pashler, H., & Harris, C. R. (2012). Is the replicability crisis overblown? Three arguments examined. *Perspectives on Psychological Science, 7*(6), 531–536. http://dx.doi.org/10.1177/1745691612463401.

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17–59). San Diego, CA: Academic Press.

Pautasso, M. (2010). Worsening file-drawer problem in the abstracts of natural, medical and social science databases. *Scientometrics.* http://dx.doi.org/10.1007/s11192-010-0233-5.

Rasinski, K. A., Visser, P. S., Zagatsky, M., & Rickett, E. M. (2005). Using implicit goal priming to improve the quality of self-report data. *Journal of Experimental Social Psychology, 41*(3), 321–327.

Rosenthal, R. (1979). An introduction to the file drawer problem. *Psychological Bulletin, 86*, 638–641.

Schvaneveldt, R. W., & McDonald, J. E. (1981). Semantic context and the encoding of words: Evidence for two modes of stimulus analysis. *Journal of Experimental Psychology. Human Perception and Performance, 7*, 673–687.

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.

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.

Srivastava, S. (2012). *What counts as a successful or failed replication?* ([October 5, 2012] Retrieved November 19, 2012, from http://hardsci.wordpress.com/).

Sutton, A. J., Duval, S. J., Tweedie, R. L., Abrams, K. R., & Jones, D. R. (2000). Empirical assessment of effect of publication bias on meta-analysis. *BMJ, 320*, 1574.

Yong, E. (2012a). Nobel laureate challenges psychologists to clean up their act. *Nature* http://www.nature.com/news/nobel-laureate-challenges-psychologists-to-clean-up-their-act-1.11535.

Yong, E. (2012b). *A failed replication attempt draws a scathing personal attack from a psychology professor.* Discover Magazine.

Young, N. S., Ioannidis, J. P. A., & Al-Ubaydli, O. (2008). Why current publication practices may distort science. *PLoS Medicine, 5*, 1418–1422.

---

[5] One might have hoped that meta-analysis could shed light on this question, but there is little reason to think that this approach can circumvent publication bias (Ferguson & Heene, 2012; Sutton, Duval, Tweedie, Abrams, & Jones, 2000).