The Effects of Implicit Religious Primes on Dictator Game Allocations: A Preregistered Replication Experiment

AQ:au, 1

Cristina M. Gomes and Michael E. McCullough
University of Miami

Shariff and Norenzayan (2007) discovered that people allocate more money to anonymous strangers in a dictator game following a scrambled sentence task that involved words with religious meanings. We conducted a direct replication of key elements of Shariff and Norenzayan's (2007) Experiment 2, with some additional changes. Specifically, we (a) collected data from a much larger sample of participants ($N = 650$); (b) added a second religious priming condition that attempted to prime thoughts of religion less conspicuously; (c) modified the wording of some of their task explanations to avoid deceiving our participants; (d) added a more explicit awareness probe; (e) reduced prime-probe time; and (f) performed statistical analyses that are more appropriate for non-normal data. We did not find a statistically significant effect for religious priming. Additional tests for possible between-subjects moderators of the religious priming effect also yielded nonsignificant results. A small-scale meta-analysis, which included all known studies investigating the effect of religious priming on dictator game offers, suggested that the mean effect size is not different from zero, although the wide confidence intervals indicate that conclusions regarding this effect should be drawn with caution. Finally, we found some evidence of small-study effects: Studies with larger samples tended to produce smaller effects (a pattern consistent with publication bias). Overall, these results suggest that the effects of religious priming on dictator game allocations might be either not reliable or else quite sensitive to differences in methods or in the populations in which the effect has been examined.

Keywords: implicit priming, generosity, religious priming, dictator game, experimental economics

Supplemental materials: http://dx.doi.org/10.1037/xge0000027.supp

Theorists have suggested that the perceived presence of religious agents (e.g., Gods, spirits, dead ancestors) promotes prosocial behaviors, and might have thereby facilitated the development of large-scale human societies (e.g., Irons, 1991; Norenzayan & Shariff, 2008; Sosis & Alcorta, 2003). Although many researchers have found relations between measures of religiousness and measures of prosociality (e.g., Batson et al., 1989; Sosis & Ruffle, 2004), until quite recently most of these findings were correlational and therefore did not allow researchers to establish the direction of causality. To surmount this limitation, Shariff and Norenzayan (2007) developed a novel experimental approach that involved implicitly activating God concepts with a scrambled sentence task (Srull & Wyer, 1979), which has been used in previous research to prime concepts such as rudeness and old age (Bargh, Chen, & Burrows, 1996), personal values (Verplanken &

Cristina M. Gomes and Michael E. McCullough, Department of Psychology, University of Miami.

AQ: 10   Correspondence concerning this article should be addressed to Cristina M. Gomes or Michael E. McCullough, Department of Psychology, University of Miami, PO Box 248185, Miami, FL 33124. E-mail: c.gomes@miami.edu or mikem@miami.edu

Holland, 2002), and thoughts about money (Vohs, Mead, & Goode, 2006).

To investigate the relation between religion and generosity, Shariff and Norenzayan (2007, Experiment 1) implicitly primed half of their participants with religious concepts (the other half was not primed) before they participated in a dictator game in which participants were endowed with a sum of money that they could then split with an anonymous (sham) recipient. Shariff and Norenzayan (2007) found that religiously primed participants allocated more money ($4.56, on average) to anonymous strangers than did participants who were not primed, $t(48) = 3.73$, $p = .001$, mean difference (MD) = $2.40, 95% CI [$1.10, $3.69], effect size $g = 1.03$ (recalculated based on published data). In Experiment 2, which was a replication and extension of Experiment 1, Shariff and Norenzayan (2007) compared the dictator game allocations of participants who had been primed with God concepts to those of participants who either (a) had been primed with secular concepts associated with morality or (b) completed a neutral-prime control task. Participants primed with religious or secular concepts allocated more money to anonymous strangers than did those who completed a control task (Religious prime: $t(48) = 2.47$, $p = .01$, MD = $2.0, 95% CI [$0.37, $3.63], $d = 0.69$; Secular Prime: $t(48) = 2.29$, $p = .03$, MD = $1.88, 95% CI [$0.23, $3.53], $g = 0.69$; recalculated based on published data). Shariff and Norenzayan (2007) interpreted their results as evidence that the perceived presence of God or moral institutions, activated through implicit priming, increases pro-social behavior.

This influential study, which according to Google Scholar has been cited over 500 times, has been widely emulated and has stimulated a great deal of empirical and theoretical innovation in the scientific study of religion, even though concerns about the meaning of the results have appeared in print (e.g., Randolph-Seng & Nielsen, 2008), and even though there is no preregistered direct replication of the experiment in the published literature. One important issue is that both of Shariff and Norenzayan's (2007) experiments had relatively small samples ($n = 25$ in each cell of a two-group design for their Experiment 1, and $n = 25$ in each cell of a three-group design for Experiment 2), which could account for the somewhat low statistical power of their experiments (the main effect in Experiment 1 had a power $= 0.54$, the main effect in Experiment 2 had a power $= 0.52$; calculated based on their published data). This is a potentially serious problem because studies with small sample sizes and low statistical power (and thus, wide confidence intervals for their effect size estimates) can lead to large overestimates of effect size and low reproducibility of results. Furthermore, Shariff and Norenzayan (2007) analyzed their data via multiple t tests, which are not well-suited to the distributional properties of dictator game offers (typically zero-inflated, with a large lump of participants who transfer 50% of their endowment, and a smaller but still substantial lump of participants who transfer 100% of their endowment; Engel, 2011). The use of t tests on data that are not normally distributed can also inflate p values above their nominal levels, thereby leading to Type I errors.

Finally, implicit priming methods in general have been criticized by several authors who have questioned the validity and interpretation of the results (e.g., Bower, 2012; Doyen, Klein, Pichon, & Cleeremans, 2012; Doyen, Klein, Simons, & Cleeremans, 2014; Yong, 2012). In particular, the psychological processes underlying implicit priming effects are not well understood. Social psychologists, for instance, have frequently used supposedly implicit stimuli as a way to investigate how primes influence behavior automatically (i.e., unconsciously), and Shariff and Norenzayan (2007) stated that their technique was novel in part because the experimental procedure "activates God concepts implicitly, without having participants consciously reflect on these concepts" (p. 804). However, many experiments that have used implicit priming methods such as the Scrambled Sentence Task—including Shariff and Norenzayan's (2007) experiments—have not fully tested whether participants were consciously aware of the primed concepts (Doyen et al., 2014).

Conscious awareness of primes can perhaps be avoided more effectively in scrambled sentence tasks by embedding each word that is related to the primed concept within a sentence that itself is unrelated to the primed concept (Bargh & Chartrand, 2000; Randolph-Seng & Nielsen, 2008). Unfortunately, many of Shariff and Norenzayan's (2007) religious prime words were set within religious sentences (e.g., the prime "God" within the sentence "give thanks to God"). To confirm that the priming effects documented by Shariff and Norenzayan (2007) can indeed be obtained without conscious awareness, it would be useful to compare the dictator game offers of participants who are primed with religious words embedded in religious sentences to the dictator game offers of participants who are primed with the same religious words embedded in nonreligious sentences.

Another way to investigate whether priming is conscious or unconscious is to probe participants by asking questions related to the primed concept. For a prime-awareness probe method to be successful, however, it should directly inquire about the primed concept (in this case, religion) and should be administered immediately after the measurement of the dependent variable (Bargh & Chartrand, 2000; Doyen et al., 2014; Randolph-Seng & Nielsen, 2008). If the time between the prime and the awareness probe is too long, the probe might be measuring memory for the content of the primes rather than conscious awareness of the content of the primes (Dixon, 1981; Doyen et al., 2014). Shariff and Norenzayan (2007) did not explicitly ask participants whether they recalled seeing religious words, but instead asked them whether they knew what the study was about, which is a different type of prime probe. Also, the probe was delivered at the end of the study, after participants had completed several questionnaires. Therefore, it is unclear whether participants in Shariff and Norenzayan's (2007) study were in fact unaware of the concept that Shariff and Norenzayan (2007) attempted to prime.[1]

Other Relevant Experiments

A few efforts to replicate and extend Shariff and Norenzayan's (2007) results to other samples or to other measures of generosity have been conducted, though none of them attempted to replicate Shariff and Norenzayan's (2007) methods exactly, and none of them had preregistered hypotheses or descriptions of methods. First, in an unpublished thesis, Hurst (2014) essentially replicated Shariff and Norenzayan's (2007) main finding: Participants primed with God concepts allocated more money to anonymous strangers than did participants primed with neutral concepts: $F(2, 116) = 5.84$, $p = .004$. However, Hurst's experiment was conducted online (whereas Shariff and Norenzayan's (2007) experiments were conducted via face-to-face interaction with experimenters), and it is not clear that Hurst used the same wording and task explanations as did Shariff and Norenzayan (2007).

Second, Ahmed and Salas (2011) also successfully replicated Shariff and Norenzayan's (2007) main result, $t(222) = 3.29$, $p = .001$ on university students in Chile, but using different scrambled sentences and different task explanations than did Shariff and Norenzayan (2007). Benjamin, Choi, and Fisher (2010) used similar methods to test the effect of religious priming on generosity in students at Cornell University, using $1 stakes. However, these researchers were unable to replicate Shariff and Norenzayan's (2007) main result. Horton, Rand, and Zeckhauser (2011) and Rand et al. (2014) used the Amazon Mechanical Turk population to test the effect of religious priming on the Prisoners Dilemma game (PD) rather than the dictator game, using $1 stakes. Both of these latter studies showed that participants primed with religious concepts were significantly more likely to cooperate with another participant (i.e., choose the *cooperate* option rather than the *defect* option in a PD, which causes the cooperator's partner to receive

AQ: 2

Fn1

---

[1] It should be noted, however, that in other studies that used the same religious priming method (i.e., scrambled sentence tasks) in conjunction with funnel debriefing procedures (Bargh & Chartrand, 2000), participants did not recognize the semantic content of the priming task or the relationship between the priming task and the dependent variable (e.g., Laurin, Kay, & Fitzsimons, 2012).

income while simultaneously causing the cooperator to also receive income, although less than he or she would have by playing the *defect option*) than participants not primed with religious concepts. However, this effect was significantly only for participants with certain religious beliefs. Other researchers have studied the effect of religious priming on generosity in other types of tasks, such as volunteering for an environmental cause (Sasaki et al., 2013), donating to organizations assisting victims of contagious disease (Preston & Ritter, 2013), and willingness to take charity pamphlets (Pichon, Boccato, & Saroglou, 2007; see Shariff, Willard, Andersen, & Norenzayan, 2015, for a review and meta-analysis). Most of these studies have found significant results, suggesting that implicit religious primes can influence other measures of generosity. However, as mentioned above, none of the studies was a direct replication of Shariff and Norenzayan (2007), and none had preregistered hypotheses or methods. As a result, they cannot necessarily be considered a completely reliable guide to the replicability of Shariff and Norenzayan's (2007) basic finding. Nevertheless, the fact that most of these experiments appear to reproduce the same basic finding gives the impression that the general effect of religious priming on generosity is reliable and robust to methodological variations. It is for this very reason that a large-scale, preregistered, methodologically "close" replication of Shariff and Norenzayan's (2007) experiment is called for.

In our experiment, we aimed to replicate Shariff and Norenzayan's (2007) Experiment 2, using the same priming materials and dependent variable they used, while also addressing some of the methodological questions we described above. Because our goal was not to compare the secular and religious priming conditions, but, rather, to determine whether religious priming increases generosity relative to a control condition, we did not run participants in their secular priming condition.

We made several modifications that enabled us to study the nature and robustness of Shariff and Norenzayan's (2007) original results more definitively. First, we collected data from 650 participants in three conditions—approximately 200 per cell—which is substantially larger than the 25 participants per cell that Shariff and Norenzayan (2007) collected for their Experiment 2. We recruited participants from two different sites: (a) students, staff, and visitors at the University of Miami (UM); and (b) visitors to an office of the Florida Department of Highway and Motor Vehicles (DMV) in Miami. Our larger sample size allowed us to estimate the effects of religious primes on dictator game allocations with higher statistical precision.

Second, we added a second religious priming condition, which we called the enhanced implicit priming condition (hereafter, ERP), using the same religious words as in the Shariff and Norenzayan's (2007) standard implicit religious priming condition (hereafter, SRP), but with the religious words embedded in nonreligious sentences. Thus, in the enhanced implicit religious priming condition, we varied the context of the sentences in hopes of making the primes more inconspicuous. This modification enabled us to investigate whether Shariff and Norenzayan's (2007) standard implicit religious prime (i.e., one that uses religious words in religious contexts) has a different effect on participants' dictator game allocations than does an enhanced religious implicit prime (i.e., one that uses religious words in nonreligious contexts). In addition to the two religious priming conditions, we also included

Shariff and Norenzayan's (2007) control condition, which involves scrambled sentences that do not include any religious words.

Third, we added an explicit awareness probe in addition to the more open-ended awareness probe that Shariff and Norenzayan (2007) used: Specifically, we asked participants whether they remembered having seen religious words to directly determine whether they were aware of the primed concept at the time they made their dictator game allocations. Fourth, we reduced prime-probe time by introducing the probes directly after measurement of the dependent variable (i.e., the dictator game) instead of after the final study questionnaire.

Finally, in addition to using *t* tests, we used Kruskal-Wallis tests, which test for differences in the medians of two or three different distributions and only assume ordered categorical data (i.e., these statistical tests make no distributional assumptions). We also used generalized linear models (GLM) to investigate the effect of the independent variables (condition, religiosity) on our dependent variable (number of dollars shared), while controlling for other variables (e.g., sex, age), and using an error distribution that is more appropriate for the allocations typically obtained in dictator games (Engel, 2011).

Method

Participants

We collected data on 650 subjects[2]: 358 from the UM (students, staff, and visitors) and 292 from visitors to an office of the DMV in Miami. UM participants were recruited at several locations on the UM's Coral Gables campus. These locations included the Patti and Allan Herbert Wellness Center and the breezeway in front of the Otto G. Richter Library. All adults who approached the recruitment booths were eligible for inclusion. Therefore, our UM sample included students, staff, faculty, and visitors to the university. DMV participants were recruited at the DMV office located inside the Mall of the Americas in Miami, Florida.

Procedures

Ara Norenzayan and Azim Shariff kindly shared the materials that they used for their religious priming condition and control condition in Experiment 2, and we used those materials as precisely as possible, with the exception of a few changes in wording to (a) avoid having to deceive our participants unnecessarily and (b) clarify some of the instructions. For example, in our instructions for the priming materials, we modified the sentence "Please complete the following verbal fluency task" to "Please complete the following task," the sentence "You have been chosen as the giver in this economic task. The next participant will be the receiver" to "For this task you will be the giver and another

---

[2] We initially based our sample size on the number of participants necessary to obtain a power of at least 0.9, given the effect size and *p* value of Experiment 2 of Shariff and Norenzayan (2007). We determined that we would need 50 subjects per condition. Given that it is unrealistic to expect exactly the same effect size when studying different populations, we then decided to double this sample size, which would result in a minimum of 600 subjects (100 * 3 conditions * 2 locations). An experiment with this sample size, and the effect size and *p* value reported by Shariff and Norenzayan (2007) for Experiment 2, would have a power >.99.

4   GOMES AND McCULLOUGH

participant will be the receiver," and made other minor changes as well (see online supplemental material for an exhaustive description of all methodological differences). Finally, we used different wording than the original authors used to assure giver–experimenter anonymity; however, we made it clear in the recruitment material, in the consent form, and in the scripted verbal explanation that all their responses would be anonymous and would remain confidential throughout. All other instructions, the priming sentences, and the game procedures were the same as those that Shariff and Norenzayan (2007) used for their Experiment 2. We had anticipated that many of our participants would not be fluent English speakers, so we had a Spanish version of the study materials translated and then back-translated by a certified English–Spanish translator and invited participants to complete the study in the language they preferred.

Once the experiment began, participants were seated in a private booth for its entire duration. On the table in front of them, participants found a computer, an envelope with $10, and a locked box. Before beginning the study, participants chose whether to complete the experiment in English or in Spanish. Next, participants were instructed that their participation was anonymous and that none of their answers or decisions would be shared with anyone, including other participants. We followed Shariff and Norenzayan's (2007) procedure for insuring anonymity: participants were alone and behind closed doors for the entire duration of the experiment; they did not enter any identifying information into the questionnaire; and they put the envelope with the money they left behind for the recipient into the slot of a locked box (as done in voting polls). To our knowledge, the only substantive difference between both experiments regarding anonymity assurance is that Shariff and Norenzayan (2007) used deception, and we did not (i.e., in their verbal explanation they told participants that the box would be given directly to the receiver whereas we did not say anything regarding the final destination of the box; we also labeled the word task differently and we used less explicit phrases when assuring anonymity to our participants).

Then the experimenter asked participants to read the instructions on the computer, which prompted them to read the Informed Consent Form and indicate their consent. Next, participants took part in a scrambled sentence task (Srull & Wyer, 1979) that required them to unscramble 10 five-word sentences, dropping an irrelevant word from each sentence to create a grammatically correct four-word sentence (see Appendix A). Each participant was randomly assigned to one of three conditions: control condition, standard religious prime (SRP) condition, or enhanced religious prime (ERP) condition. In the control condition, none of the 10 sentences had religious words; thus, participants were presented with 10 sentences with neutral words (neutral prime control). In the SRP, 5 of the 10 sentences had one religious word (spirit, divine, God, sacred, and prophet; same sentences used in Shariff & Norenzayan, 2007), and four nonreligious words, which together made a religious sentence in four of the five sentences (the fifth sentence is nonreligious; see Appendix). The remaining five sentences were identical to five of the sentences used in the control condition. In the ERP condition, we used five different sentences than Shariff and Norenzayan (2007) used for the religious words in order to change the context of the sentences from religious to nonreligious. In this condition, 5 of the 10 sentences contained one religious word (spirit, divine, God, sacred and prophet; same

words used in Shariff & Norenzayan, 2007) and four nonreligious words—different from the ones used in Shariff and Norenzayan (2007)—that together formed nonreligious sentences. The remaining five sentences had the same five nonreligious words used in the SRP condition and the control condition.

After the priming manipulation, participants were asked via computerized instructions to open the envelope in front of them, which contained 10 $1 bills. They were then told that they could choose to keep all of that money for themselves or leave some or all of the money for a participant who would do the experiment in the future. This procedure achieved participants' anonymity by ensuring participants that they would never come into contact with the receiver, who would therefore, never know their identity.

After participants decided how much money to leave for the future participant, they were prompted to answer a series of awareness probe questions that enabled us to assess their understanding of the study, their understanding of the scrambled sentence task, whether they recalled any particular theme in the sentences, and whether they recalled having seen religious words in the sentences.

Next, participants completed a short demographic questionnaire for gathering information on their age, sex, annual income, religiosity (religious, spiritual, agnostic, or atheist), religious affiliation (Christian, Muslim, Jewish, Buddhist, Hindu, Other), race, and degree of belief in God (on a Likert-type scale with response options ranging from 1 to 7).

After participants completed the experiment, they were debriefed. We then announced that they were designated to be the recipients of a previous dictator's allocation decision and so would also receive the money the previous participant left behind for his or her respective recipient. After receiving this additional money, participants were given a chance to ask further questions about the experiment and were dismissed.

Data Analyses

We used t tests to replicate Shariff and Norenzayan's (2007) analyses. Additionally, we used nonparametric Kruskal-Wallis tests to compare median dictator game offers in the three conditions and GLMs to investigate (as in Shariff & Norenzayan, 2007) whether condition (control, SRP, and ERP), religiosity, and select demographic variables (e.g., sex, age, income) affected the amount of money participants allocated to their respective dictator game recipients. GLMs can include several fixed effects (i.e., predictor variables that influence the mean) in a single model and are generally superior to methods commonly used to investigate the effect of different experimental conditions on a dependent variable when the response variable is not normally distributed (i.e., the GLMs model the error structure correctly; Bolker, 2007).

Our main analyses involved building two GLMs with all of the variables that Shariff and Norenzayan (2007) examined in Experiment 2. Thus, the first of these two GLMs included amount sent as a dependent variable, and condition (control, SRP, and ERP), religiosity (atheists vs. theists), and an interaction between condition and religiosity as independent variables. The second GLM used a measure of belief in God (a continuous variable) as a substitute for the atheists-versus-theists measure of religiosity. These two analyses map exactly on to the main analyses that Shariff and Norenzayan (2007) conducted to test their hypotheses in their Experiment 2.

We did additional analyses, on a purely exploratory basis, to investigate whether other religious, demographic, or methodological variables moderated the effects of our religious primes. To do so, we first built a main effects model that included the following independent variables: recruitment site (university vs. DMV), religious affiliation (agnostic, atheist, Buddhist, Christian, Hindu, Jewish, Muslim, spiritual without an organized religion, other), sex (male vs. female), age category (broken into three categories that we hoped would approximate demographically meaningful cut points: 18–25, 26–40, 41 or older), annual income (broken into seven groups of $10,000 increments each), and a dummy variable for whether participants were of Hispanic heritage (yes vs. no). We then compared the main effects model to a series of interaction models each of which resulted from adding to the main effects model an interaction between condition and each of the independent variables mentioned above. Study language was correlated with annual income: participants who chose to complete the experiment in Spanish had significantly lower annual incomes than those who did the study in English, $b = 0.34$, $p < .0001$. Thus, we did a separate analysis substituting annual income for study language.

Additionally, we investigated whether prime awareness moderated the effect of our religious primes by including two of our prime probes as independent variables in two additional models. The first included the yes/no prime probe exploring whether participants recalled any particular theme in the sentences, and the interaction between the probe and condition; and the second model included the yes/no prime probe exploring whether participants recalled having seen religious words in the sentences, and the interaction between the probe and condition.[3] In all models, we measured the effect of each interaction and categorical variable separately by doing an F test comparing the fit of the full model (i.e., one with all the above independent variables plus one interaction) to the reduced model (i.e., one with all the above independent variables but without the interaction; Bolker, 2007) using the analysis of variance (ANOVA) function in R.

Finally, at a reviewer's request, we performed a meta-analysis to estimate the effect of religious priming on dictator game generosity using all of the known experiments that have been conducted on this topic. All analyses were done in R (version 3.1.1) using the stats package for the t tests (t.test function) and Kruskal-Wallis test (kruskal.test function), the base package for GLMs (glm function), and the metafor package for the meta-analysis (rma function).

Pre-Registration

A summary of these methods was preregistered in the Open Science Framework on March 11, 2014, and can be found at https://osf.io/i53qz/registrations/. A draft of the introduction and methods sections of this paper was uploaded on July 27, 2014, and can be found at https://osf.io/i53qz/. Data collection started on October 14, 2013, a few months prior to registration. Although 170 (26%) of the 650 participants had already taken part in the study when it was registered, our protocol, described in the preregistration and uploaded draft, remained the same throughout the entire data collection period. No analyses were done until data collection was completed.

To verify that the data collected prior to registration did not differ from the data collected after registration, we added to the two main models (the one with religiosity and the other with the 7-point measure of belief in God) a binary variable that indicated whether subjects' data were collected before or after the preregistration process. In addition, we added the interactions between condition, religiosity, and our registration variable. Unsurprisingly, none of these variables were significant, indicating that the data collected prior to preregistration process did not differ from the data collected after. Furthermore, religious priming had no significant effect on dictator game offers in either the data collected prior to registration or in the data collected after registration. Thus, we are confident that the flaws in our preregistration process did not compromise the conclusions we attempted to draw from the other analyses reported here.

Results

Population Characteristics

Of the 650 participants recruited, 56% were males and 44% were females, and 51% were Hispanic whereas the remaining 49% were not. On average, participants were 30 years of age, ranging between 18 and 85 years old (354 participants were 18 to 25 years; 170 participants were 26 to 40 years; 126 participants were 41+ years). A total of 429 participants indicated an identification with a religion (68% Christian, 7% Jewish, 3% Hindu, 3% Buddhist, 2% Muslim, 17% other) whereas 220 identified as either spiritual without a religion, agnostic, or atheist (one person did not answer this question). Following Shariff and Norenzayan (2007), participants were considered atheist if, in addition to identifying as atheist or agnostic, they scored below the midpoint of the 7-point scale on the belief in God question. Participants who did not indicate a religious organization but scored higher than the midpoint on the belief in God question were considered theist, along with those who did state specific religious identifications. These group assignments resulted in 115 participants identified as atheists and 534 participants identified as theists. About 15% of our participants completed the Spanish-language version of the experiment, whereas the remaining 85% completed the English-language version.

Because of a programming error, random assignment to conditions was performed with replacement rather than without replacement, which resulted in a slightly unbalanced sample (N: Control = 216, SRP= 239, ERP = 195). However, a chi-square analysis that compared the observed sample sizes to the expected balanced sample sizes revealed no significant difference, $p = .12$, suggesting (as one would expect) that the data were nevertheless sampled from the expected uniform distribution.

Suspicion Probe

When we asked participants the open-ended question designed to determine whether they could identify the purpose of the experiment, only 9 out of 650 participants related the religious theme in the scrambled sentence task to generosity in the dictator game. All 9 of those subjects had participated in the SRP condition.

---

[3] A third prime probe exploring whether participants knew the hypothesis of the study did not have enough variability across conditions for it to converge. We therefore did not include it in this model.

Fn3

Those 9 participants were excluded from all analyses; nevertheless, including these data did not significantly change the results. In response to the open-ended question regarding whether the scrambled sentences had had a common theme, 104 participants (44%) in the SRP condition versus 14 participants (7%) in the ERP condition reported having noticed a religious theme, chi-square test: $p < .0001$. Finally, when specifically asked if they recalled having seen religious words in the sentences, 222 (93%) and 170 (87%) participants in the standard and ERP conditions, respectively, reported having seen religious words, chi-square test: $p < .01$. Aside from the 9 participants mentioned above, no additional participants connected the religious theme to the dictator game. Therefore, no other participants were excluded from the analyses.

Effect of Primes and Religiosity

Participants in the control condition left, on average, $4.49 ($SD = 3.49$) for other participants, with 28% leaving $1 or less, 32% leaving $5, and 27% leaving more than $5. Participants in the SRP condition left, on average, $4.28 ($SD = 3.67$) for other participants, with 34% leaving $1 or less, 26% leaving $5, and 27% leaving more than $5. Finally, participants in the ERP condition left, on average, $4.7 ($SD = 3.72$) for other participants, with 29% leaving $1 or less, 24% leaving $5, and 31% leaving more than $5 (Table 1, Figure 1).

T test analyses showed that the differences between the control condition and the two religious conditions, and between the SRP condition and the ERP condition, were not statistically significant (control vs. SRP: $t(214) = 0.63$, $p = .53$, mean difference (MD) = $0.21, 95% CI [–$0.45, $0.88], $g = -0.059$; control versus ERP: $t(228) = -0.58$, $p = .56$, MD = 0.21, 95% CI [–$0.91, $0.49], $g = 0.058$); SRP versus ERP: $t(193) = -1.17$, $p = .24$, MD = 0.42, 95% CI [–$1.13, $0.29], $g = 0.11$). A Kruskal-Wallis test that compared the median transfers across the three conditions was also nonsignificant, $\chi^2 = 1.64$, $df = 2$, $N = 641$, $p = .44$. In the GLM with a quasi-Poisson error structure (we did not use a Poisson model because diagnostics showed that the distribution of dictator game offers was overdispersed; Bolker, 2007), none of the variables significantly predicted how much money participants sent (see Table 2): condition (95% CI [−0.19, 0.17]); religiosity (95% CI [−0.37, 0.13]); condition*religiosity (95% CI [−0.15, 0.24]).[4] A model that included the 7-point self-report measure of belief in God instead of the religiosity measure also revealed a nonsignificant relationship between amount sent and condition (see Table 2). Finally, none of the interactions resulting from combining condition with recruitment site, religious affiliation, age, sex, Hispanic heritage, annual income, study language, or prime probe had a significant effect on amount sent. In no case did
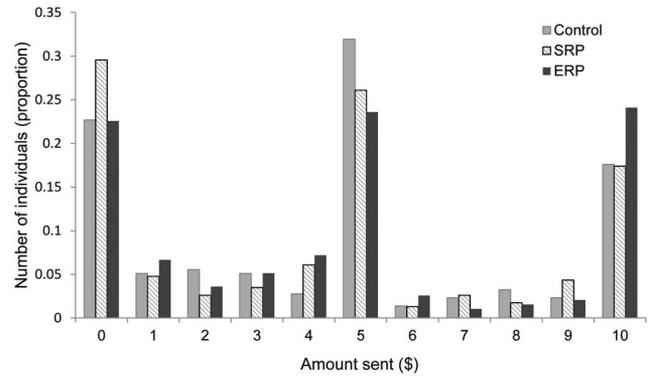
Figure 1. Frequency distribution (in proportion, i.e., number of individuals per total number of participants in each condition) of money offered in the control condition, the standard implicit prime (SRP) condition, and the enhanced implicit religious prime (ERP) condition.

controlling for the effect of the above independent variables and the resulting interactions modify the nonsignificant effect of religious priming on generosity. However, sex, annual income, and study language had main effects on amount sent: women shared more money than men, $F = 6.54$, $p = .01$, higher-income participants shared more money than did lower-income participants, $b = 0.02 \pm 0.01$, $p = .03$, and participants who chose to complete the experiment in English shared more money than those who chose to complete it in Spanish, $F = 8.09$, $p = .004$.

Meta-Analysis

To compile data for our meta-analysis, we searched for additional experiments that statistically compared the effect of religious priming on generosity in dictator games and performed a random-effects meta-analysis upon them. In addition to Shariff and Norenzayan's (2007) experiments and the present experiment, we found three experiments that met these criteria, totaling six studies (see Table 3). To statistically combine these experiments, we used the Hedges g statistic to express the standardized mean difference between the religious prime condition and the control condition in dictator game allocations.[5] To represent the data from the experiment we present here, we used the effect size for the comparison of the standard implicit religious prime condition to the control condition (i.e., the data for the enhanced implicit religious prime condition was not included in this meta-analysis). We did not include data from other experiments (e.g., those reviewed in Shariff et al. 2015) that could be conceptualized as measures of prosociality because they clearly are attempts to estimate different parameters: Our goal here is to know whether religious priming

Table 1

Descriptive Statistics of the Control Condition, the SRP Condition, and the ERP Condition

Parameter	Control	SRP	ERP
N	216	230	195
Median	5	5	5
Mean	4.49	4.28	4.7
SD	3.49	3.67	3.72

---

[4] At the request of a reviewer, we reanalyzed Shariff and Norenzayan's (2007) study using the data they reported in Figure 2, and following the same statistical procedure described here. A Kruskal-Wallis analysis showed a significant relationship between amount sent (\$) and condition, $\chi^2 = 6.35$, $df = 1$, $N = 50$, $p = .012$. This result was further confirmed using a GLM with a quasi-Poisson error structure, GLM: $b \pm se = 0.58 \pm 0.24$, $p = .02$, 95% CI [0.11, 1.07], $g = 1.03$.

[5] We added the Knapp and Hartung (2003) adjustment recommended for small meta-analytic samples (IntHout, Ioannidis, & Borm, 2014).

Table 2

Variables Influencing the Amount of Money Participants Sent Their Partner

Model	Predictor variables	b ± se	p value
	Condition	−0.011 ± .09	0.90
1	Religiosity (atheist, theist)	0.99 (*F*)	0.31
	Condition*religiosity	0.16 (*F*)	0.68
	Condition	−0.07 ± .09	0.48
2	Religiosity (7-point scale)	−0.01 ± .02	0.57
	Condition*religiosity	1.05 (*F*)	0.31

Note. We use the beta estimate (B) to present the effect size of continuous variables and the *F* statistic to present the effect size of categorical variables and interactions.

influences dictator game generosity; not prosociality more generally.

The overall weighted effect size of religious priming on dictator game generosity across all six studies did not reach statistical significance, although it was positive in sign, $g = 0.37$, $SE = 0.18$, $p = .09$, 95% CI [−0.09, 0.83] (Figure 2). We also found evidence for publication bias, as noted by the significant Egger's test for funnel plot asymmetry (Egger, Smith, Schneider, & Minder, 1997): $b = −0.54$, $t = 4.39$, $df = 4$, $p = .01$. Tests of heterogeneity suggest that a substantial amount of the variation in effect sizes is due to between-studies differences rather than sampling error, $Q = 38.57$, $df = 5$, $p < .0001$, $I^2 = 89\%$. These heterogeneity statistics imply that the six experiments included here are estimating different parameters, perhaps due to between-studies differences in population or methodology.

Additionally, we used the precision-effect test - precision-effect estimate with standard error test (PET-PEESE) estimator (Stanley et al., 2013) as in Carter, Kofler, Foster, and McCullough (in press) and Carter and McCullough (2014), to estimate the effect size that would be expected in an experiment with infinite precision (i.e., a study in which the standard error is zero). When the intercept for the regression of effect size on standard error is significant (i.e., the null hypothesis that b_0 is equal to zero cannot be rejected), as was the case here, PET: $b_0 = −0.54$, $p = .03$; 95% CI [−0.98, −0.09], Stanley et al. (2013) suggest that the most accurate estimate of effect size comes from the regression of effect size on variance (i.e., standard error squared). In the case of our small-scale meta-

Table 3

Studies Included in the Meta-Analysis

Study	N	Hedges g	Variance	SE
Shariff and Norenzayan (2007)—S1	50	1.03	0.091	0.30
Shariff and Norenzayan (2007)—S2	50	0.69	0.085	0.29
Ahmed & Salas (2011)	224	0.44	0.018	0.13
Gomes and McCullough (this experiment)—SRP vs. Control	446	−0.06	0.0089	0.095
Benjamin et al. (2010)	554	−0.13	0.0072	0.085
Hurst (2014)	142	0.60	0.0299	0.17

Note. We calculated all measures of Benjamin et al. (2010), based on means and standard deviations provided by the authors; Hurst (2014) reported sample sizes, means, and standard deviations separately for atheists and theists, instead of total values per condition. To obtain the overall mean and standard deviation per condition of this experiment, we pooled the means and standard deviations, weighing them by sample size.
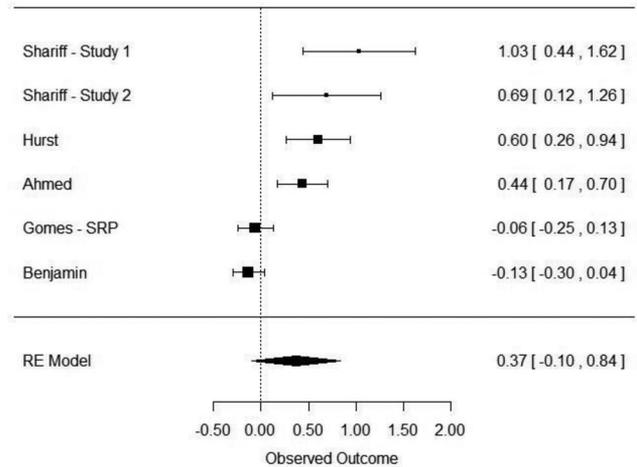
Figure 2. Forest plot showing the effect size (Hedges *g*) and confidence intervals of the overall model resulting from the random-effects meta-analysis and each one of the individual studies.

analysis, this regression gave a negative and nonsignificant estimate for the intercept, PEESE: $b_0 = −0.12$, $p = .37$, although the 95% CI [−0.45, 0.21] contains some positive values. The PET-PEESE results give a slightly different perspective from which to interpret the results of the random-effects meta-analysis: When controlling for small-study effects (i.e., the tendency for the experiments in this meta-analytic sample with smaller numbers of participants to produce larger effect sizes), religious priming does not appear to have a significant effect on generosity in an idealized experiment in which the standard error (i.e., parameter estimation error due to sampling error variance) is simulated to be zero. This result is evident in Figure 2, in which larger samples with smaller standard errors produce smaller effects. However, given the very small number of studies included in these meta-analyses—as well as the large confidence interval for the intercept in the PEESE model—we counsel extreme caution in interpreting our results: Even a small number of additional experiments could substantially change the pattern of results we discovered here (Sterne et al., 2011). These meta-analytic results should therefore be taken as a snapshot of the current literature rather than as a definitive summary of whether religious priming affects dictator game giving.

Discussion

Shariff and Norenzayan's (2007) religious priming experiments were originally designed to test the hypothesis that the perceived presence of religious agents promotes prosociality. In the present experiment, which included two large samples of participants from Miami-Dade County, we were unable to reproduce the positive significant effect of religious priming on generosity in dictator games that Shariff and Norenzayan (2007) discovered. Furthermore, according to our moderator analyses, religious priming did not increase dictator game offers in any of our demographic groups; that is, age, gender, annual income, Hispanic heritage, religious affiliation, experiment location and experiment language did not moderate the effect of religious priming on generosity in this large and demographically heterogeneous sample.

Our findings, which come from an effort to replicate as closely as possible (without deceiving participants about any features of the experiment) Shariff and Norenzayan's (2007) methods, suggests that the effect of religious priming on generosity may not be particularly robust to methodological variations and/or population differences among studies.

The most obvious methodological difference between Shariff and Norenzayan's (2007) experiments and ours is sample size. Consequently, our estimates were considerably more precise. Even so, other researchers have been able to reproduce Shariff and Norenzayan's (2007) results successfully using slightly different methods and larger sample sizes (e.g., Hurst, 2014, but cf. Benjamin et al., 2010), so it is unclear that statistical power alone is responsible for the differences between our results and theirs.

One noteworthy difference between Shariff and Norenzayan's (2007) results and the results presented here is that participants in our experiment gave more money in the control condition ($4.49 on average) than they did in Shariff and Norenzayan's (2007) control conditions ($2.6 on average), suggesting that our participants were more generous on average than their participants were. As one editorial reviewer observed, it is not inconceivable that our efforts to avoid deception (see Methods) reduced participants' perceptions of giver–experimenter anonymity (giver–receiver anonymity was assured in the same way as the original authors did), thereby leading to more generous dictator game offers (Hoffman, McCabe, & Smith, 1996) and consequently a failure of the religious prime. However, several facts speak against this possibility.

First, the distribution of dictator game offers in our control condition looked very similar to the typical dictator game distributions of nonstudent participants handling real money (Kolmogorov–Smirnov: $D = 0.12$, $p = .11$; recalculated from Figures 6 and 8 in Engel, 2011): our dictators transferred, on average, $4.49 ($SD = 3.49$), with 27% transferring more than $5, whereas across other dictator game experiments in which nonstudent participants handled real notes or coins, dictators transferred $4.21 ($SD = 3.21$), with 24% transferring more than $5 (recalculated from Figures 6 and 8 in Engel, 2011). Thus, the levels of generosity that our subjects evinced are not extraordinarily high, given the populations we sampled and the fact that our subjects handled real money. Relatedly, our exploratory analyses of demographic variables suggest that sex, income, and study language had significant main effects on amount sent (with women, high-income participants, and participants completing the experiment in English sharing more than men, low-income participants, and participants completing the experiment in Spanish). It is therefore plausible that the differences in baseline dictator game offers between our experiment and Shariff and Norenzayan's (2007) are a consequence of differences in the demographic composition of our study populations (e.g., if we had a greater representation of high-income generous participants than Shariff & Norenzayan, 2007, did). Thus, it is not necessary to invoke a failure to guarantee subjects' anonymity to explain our contrasting findings.

Second, a few independent experiments and at least two meta-analyses have concluded that giver–experimenter anonymity has either a small nonsignificant effect or no effect at all on dictators' offers (e.g., Bolton, Katok, & Zwick, 1998; Engel, 2011; Zhang & Ortmann, 2012). Third, other replications of Shariff and Norenzayan (2007) have also diverged from their methods for securing subjects' anonymity. Most notably, Benjamin et al. (2010) did not deceive participants at all. Also, Ahmed and Salas (2011) had participants complete the experiment surrounded by other participants, which certainly should have raised subjects' concerns about anonymity. Even so, Benjamin et al. (2010) and Ahmed and Salas (2011) both found baseline dictator game offers that were similar to those in Shariff and Norenzayan (2007), suggesting that adhering exactly to Shariff and Norenzayan's (2007) original methods for reassuring participants' anonymity is not a necessary precondition for establishing low baseline levels of generosity.

Fourth, although higher mean dictator game offers in our three conditions could in principle create a restriction of range that limited our ability to detect true differences between our experimental conditions, the standard deviations suggest otherwise: despite our higher mean transfers, the standard deviations for our three conditions (control = 3.49, SRP = 3.67, ERP = 3.72) were actually larger than those in Shariff and Norenzayan's (2007) two experiments (SD, Study 2: control = 2.69, religious priming = 3.03), allowing enough statistical room for the $2 increase associated with the religious prime. In light of these four considerations, therefore, it is unlikely that the 19% difference in baseline generosity between Shariff and Norenzayan's (2007) experiment and ours could be responsible, on either theoretical or statistical grounds, for the inconsistencies between their results and ours. Additional experiments (ideally preregistered ones comparing different methods in populations with similar demographics or vice versa) could settle this matter definitively.

Meta-Analytic Conclusions

We also performed a random-effects meta-analytic synthesis on all known experiments that estimated the effect of the Shariff and Norenzayan (2007) religious priming technique on dictator game transfers. These meta-analyses indicated that overall, the effect of religious priming on dictator game offers is probably quite small: The PET-PEESE meta-analytic results suggested that the effect size expected in an idealized study with an infinitely small standard error is small and negative rather than positive, albeit with a wide 95% CI that includes some positive values. We also found evidence consistent with publication bias inasmuch as smaller studies tended to produce larger effect size estimates. However, we wish to emphasize again that these meta-analytic results must be interpreted with extreme circumspection given the small number of studies included and the large degree of between-studies variance in parameter estimates. Nevertheless, our meta-analytic results do largely reinforce the conclusions we drew from the present experiment: The proposition that religious priming increases generosity in the dictator game evidently cannot be taken for granted.

The results we report here might seem to fit uneasily with a recent meta-analysis, suggesting that religious priming has a consistent and positive effect on a wide variety of measures of prosociality (Shariff et al., 2015). However, Shariff et al.'s (2015) meta-analysis contains exactly the same set of dictator game experiments that we included in our meta-analysis (except, of course, for the original experiment we presented herein). Had Shariff et al. (2015) meta-analyzed those results separately from the other results they examined (and had they used the same meta-analytic approach we used), they would have drawn exactly the same conclusion we have drawn here: the typical effect of religious priming on dictator game offers appears to be small, and

AQ: 6

AQ: 7

may not differ from zero—irrespectively of whatever effects it might have on other measures of prosociality or other constructs. At the very least, when we consider the results from the experiment presented here and our meta-analytic effort, we conclude that the hypothesis that religious priming manipulations generally increase dictator game offers should not be wholeheartedly accepted at this time.

How Important Is Conscious Awareness of the Concept Being Primed?

We also aimed to investigate whether being consciously aware of the primed concept could affect people's behavior in the dictator game. However, neither the SRP nor the ERP significantly affected how much money participants left for the recipient, even when considering participants' awareness of the primed concept (based on models in which we used participants responses to our prime probes as moderators of the experimental conditions). Thus, we were not able to determine whether priming participants with inconspicuous primes works better than priming participants with conspicuous ones. Nevertheless, only participants assigned to the SRP condition were able to determine the goal of the study. These participants were also much more likely to spontaneously report awareness of a religious theme within the sentences. In addition, participants in the SRP condition were slightly better at recalling (when prompted) that the task involved religious words than were participants assigned to the ERP condition. Taken collectively, these results suggest that the ERP was the more subtle "implicit prime" of religious semantic content, whereas the SRP was considerably less subtle. Experiments using primes that have significant effects on behavior will be necessary to determine the moderating effect of prime conspicuousness and conscious awareness of the primed concept on priming; nevertheless, future studies will likely benefit from using inconspicuous primes, such as those we used in our enhanced implicit prime condition, if researchers' goal is to conduct experiments in which participants' conscious awareness of the primed concept is minimized.

Conclusion

Overall, our results suggest that the effect of religious priming on generosity in the dictator game may be more limited than previously thought. If it is indeed a real phenomenon, then it appears to be sensitive to methodological and/or population differences. One might speculate, for example, that religious priming increases generosity only in populations of people who have baseline dictator offers below a particular threshold (e.g., in student samples but not in nonstudent samples; Engel, 2011). However, it is important to note that this threshold is currently unknown and is underspecified by theory.

If such differences are responsible for the discrepancy between our results and those of previous experiments, theoretical refinements and new experiments (ideally, preregistered ones) explicitly designed a priori to test those refinements will be needed. Going forward, large-scale experiments that include participants from a variety of religions, age classes, sexes, cultural groups, and income classes—with the goal of stratifying them into groups with differing baseline levels of generosity and then testing the effects of religious priming within each of those subgroups—might be particularly useful for helping us better understand whether we can confidently expect religious priming to increase generosity in the dictator game, and if so, for whom and under what conditions.

References

Ahmed, A. M., & Salas, O. (2011). Implicit influences of Christian religious representations on dictator and prisoner's dilemma game decisions. *The Journal of Socio-Economics, 40,* 242–246. http://dx.doi.org/10.1016/j.socec.2010.12.013

Bargh, J. A., & Chartrand, T. L. (2000). The mind in the middle: A practical guide to priming and automaticity research. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 253–285). New York, NY: Cambridge 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,* 230–244. http://dx.doi.org/10.1037/0022-3514.71.2.230

Batson, C. D., Oleson, K. C., Weeks, J. L., Healy, S. P., Reeves, P. J., Jennings, P., & Brown, T. (1989). Religious pro-social motivation—Is it altruistic or egoistic. *Journal of Personality and Social Psychology, 57,* 873–884. http://dx.doi.org/10.1037/0022-3514.57.5.873

Benjamin, D. J., Choi, J., & Fisher, G. (2010, April). *Religious identity and economic behavior.* NBER Working Paper No. 15925, Cambridge, England.

Bolker, B. (2007). *Ecological models and data in R.* Princeton, NJ: Princeton University Press.

Bolton, G. E., Katok, E., & Zwick, R. (1998). Dictator game giving: Rules of fairness versus acts of kindness. *International Journal of Game Theory, 27,* 269–299. http://dx.doi.org/10.1007/s001820050072

Bower, B. (2012). The hot and cold of priming: Psychologists are divided on whether unnoticed cues can influence behavior. *Science News, 181,* 26–29. http://dx.doi.org/10.1002/scin.5591811025

Carter, E., Kofler, L. M., Forster, D. E., & McCullough, M. E. (in press). A series of meta-analytic tests of the depletion effect: Self-control does not seem to rely on a limited resource. *Journal of Experimental Psychology: General.*

Carter, E. C., & McCullough, M. E. (2014). Publication bias and the limited strength model of self-control: Has the evidence for ego depletion been overestimated? *Frontiers in Psychology, 5,* 1–11. http://dx.doi.org/10.3389/fpsyg.2014.00823

Dixon, N. F. (1981). *Preconscious processing.* New York, NY: Wiley.

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

Doyen, S., Klein, O., Simons, D., & Cleeremans, A. (2014). On the other side of the mirror: Priming in cognitive social psychology. *Social Cognition, 32,* 12–32. http://dx.doi.org/10.1521/soco.2014.32.supp.12

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315,* 629–634. http://dx.doi.org/10.1136/bmj.315.7109.629

Engel, C. (2011). Dictator games: A meta study. *Experimental Economics, 14,* 583–610. http://dx.doi.org/10.1007/s10683-011-9283-7

Hoffman, E., McCabe, K., & Smith, V. L. (1996). Social distance and other-regarding behavior in dictator games. *The American Economic Review, 86,* 653–660.

Horton, J., Rand, D., & Zeckhauser, R. (2011). The online laboratory: Conducting experiments in a real labor market. *Experimental Economics, 14,* 399–425. http://dx.doi.org/10.1007/s10683-011-9273-9

Hurst, S. (2014). Religion's moral delusion (Unpublished master's thesis). London School of Economics, London, UK.

IntHout, J., Ioannidis, J. P., & Borm, G. F. (2014). The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightfor-

ward and considerably outperforms the standard DerSimonian-Laird method. *BMC Medical Research Methodology, 14,* 25. http://dx.doi.org/10.1186/1471-2288-14-25

Irons, W. (1991). How did morality evolve. *Zygon, 26,* 49–89. http://dx.doi.org/10.1111/j.1467-9744.1991.tb00803.x

Knapp, G., & Hartung, J. (2003). Improved tests for a random effects meta-regression with a single covariate. *Statistics in Medicine, 22,* 2693–2710. http://dx.doi.org/10.1002/sim.1482

Laurin, K., Kay, A. C., & Fitzsimons, G. M. (2012). Divergent effects of activating thoughts of God on self-regulation. *Journal of Personality and Social Psychology, 102,* 4–21. http://dx.doi.org/10.1037/a0025971

Norenzayan, A., & Shariff, A. F. (2008). The origin and evolution of religious prosociality. *Science, 322,* 58–62. http://dx.doi.org/10.1126/science.1158757

Pichon, I., Boccato, G., & Saroglou, V. (2007). Nonconscious influences of religion on prosociality: A priming study. *European Journal of Social Psychology, 37,* 1032–1045. http://dx.doi.org/10.1002/ejsp.416

Preston, J. L., & Ritter, R. S. (2013). Different effects of religion and God on prosociality with the ingroup and outgroup. *Personality and Social Psychology Bulletin, 39,* 1471–1483. http://dx.doi.org/10.1177/0146167213499937

Rand, D., Dreber, A., Haque, O., Kane, R., Nowak, M. A., & Coakley, S. (2014). Religious motivations for cooperation: An experimental investigation using explicit primes. *Religion.* Brain and Behavior, *4,* 31–48.

Randolph-Seng, B., & Nielsen, M. E. (2008). Is God really watching you? A response to Shariff and Norenzayan. *International Journal for the Psychology of Religion, 18,* 119–122.

Sasaki, J. Y., Kim, H. S., Mojaverian, T., Kelley, L. D. S., Park, I. Y., & Janusonis, S. (2013). Religion priming differentially increases prosocial behavior among variants of the dopamine D4 receptor (DRD4) gene. *Social Cognitive and Affective Neuroscience, 8,* 209–215. http://dx.doi.org/10.1093/scan/nsr089

Shariff, A. F., & Norenzayan, A. (2007). God is watching you: Priming God concepts increases prosocial behavior in an anonymous economic game. *Psychological Science, 18,* 803–809. http://dx.doi.org/10.1111/j.1467-9280.2007.01983.x

Shariff, A. F., Willard, A. K., Andersen, T., & Norenzayan, A. (2015). Religious priming: A meta analysis with focus on prosociality. *Social Psychology Review.* Advance online publication.

Sosis, R., & Alcorta, C. (2003). Signaling, solidarity, and the sacred: The evolution of religious behavior. *Evolutionary Anthropology, 12,* 264–274. http://dx.doi.org/10.1002/evan.10120

Sosis, R., & Ruffle, B. J. (2004). Ideology, religion, and the evolution of cooperation: Field experiments on Israeli Kibbutzim. *Socioeconomic Aspects of Human Behavioral Ecology, 23,* 89–117. http://dx.doi.org/10.1016/S0190-1281(04)23004-9

Srull, T. K., & Wyer, R. S. (1979). Role of category accessibility in the interpretation of information about persons—Some determinants and implications. *Journal of Personality and Social Psychology, 37,* 1660–1672. http://dx.doi.org/10.1037/0022-3514.37.10.1660

Stanley, T. D., Doucouliagos, H., Giles, M., Heckemeyer, J. H., Johnston, R. J., Laroche, P., . . . Rost, K. (2013). Meta-analysis of economics research reporting guidelines. *Journal of Economic Surveys, 27,* 390–394. http://dx.doi.org/10.1111/joes.12008

Sterne, J. A. C., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., . . . Higgins, J. P. T. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *British Medical Journal, 343,* d4002. http://dx.doi.org/10.1136/bmj.d4002

Verplanken, B., & Holland, R. W. (2002). Motivated decision making: Effects of activation and self-centrality of values on choices and behavior. *Journal of Personality and Social Psychology, 82,* 434–447. http://dx.doi.org/10.1037/0022-3514.82.3.434

Vohs, K. D., Mead, N. L., & Goode, M. R. (2006). The psychological consequences of money. *Science, 314,* 1154–1156. http://dx.doi.org/10.1126/science.1132491

Yong, E. (2012). Replication studies. *Nature, 485,* 298–300. http://dx.doi.org/10.1038/485298a

Zhang, L., & Ortmann, A. (2012). *A reproduction and replication of Engel's meta-study of dictator game experiments* (Research Paper No. 2012 ECON 44). UNSW Australian School of Business.

(*Appendix follows*)

Appendix A

Scrambled Sentence Task Sentences

Control Condition

1. fall was worried she always

2. shoes give replace old the

3. appreciated presence was imagine her

4. more paper it once do

5. send I over it mailed

6. saw hammer he the train

7. yesterday it finished track he

8. sky the seamless blue is

9. predictable he shoes his tied

10. prepared somewhat I was retired

Standard Religious Prime Condition (Same Sentences as Shariff & Norenzayan, 2007)

1. felt she eradicate spirit the

2. dessert divine was fork the

3. appreciated presence was imagine her

4. more paper it once do

5. send I over it mailed

6. evil thanks give god to

7. yesterday it finished track he

8. sacred was book refer the

9. reveal the future simple prophets

10. prepared somewhat I was retired

Enhanced Religious Prime Condition

1. she played spirit had high

2. dessert divine was fork the

3. appreciated presence was imagine her

4. more paper it once do

5. send I over it mailed

6. god I jog was tired

7. yesterday it finished track he

8. buildings they Europe toured sacred

9. was prophet become a Shakespeare

10. prepared somewhat I was retired