Does intuitive mindset influence belief in God? A registered replication of Shenhav, Rand and Greene (2012)

S. Adil Saribay[*]   Onurcan Yilmaz[†]   Gülay Gözde Körpe[‡]

Abstract

In 2012, two independent groups simultaneously demonstrated that intuitive mindset enhances belief in God. However, there is now some mixed evidence on both the effectiveness of manipulations used in these studies and the effect of mindset manipulation on belief in God. Thus, this proposal attempted to replicate one of those experiments (Shenhav, Rand & Greene, 2012) for the first time in a high-powered experiment using an under-represented population (Turkey). In line with the intuitive belief hypothesis, a negative correlation between reflectiveness and religious belief emerged, at least in one of the experimental conditions. In contrast to that hypothesis, however, the results revealed no effect of the cognitive style manipulation on religious belief. Although a self-report measure (Faith in Intuition) provided evidence that the manipulation worked as intended, it did not influence actual performance (Cognitive Reflection Test), suggesting a demand effect problem. Overall, the results failed to provide support for the intuitive belief hypothesis in our non-WEIRD sample, despite generally following the predicted patterns, and suggest that using stronger manipulation techniques are warranted in future studies.

Keywords: intuitive thinking, belief in God, replication, analytic cognitive style, reflection, intuition

1 Introduction

Although several approaches attempting to explain religiosity have been developed to date, there is still an ongoing debate in the literature regarding whether intuitive and reflective thinking processes have any role in religious belief or belief in God more specifically. Central to this debate, the dual-process model of the mind distinguishes between two sets of processes: Type 1 corresponds to intuitive, automated, and low-effort processes; Type 2 corresponds to analytical, controlled, and high-effort processes (Evans & Stanovich, 2013). As belief in God is also based on the relatively automatic beliefs acquired during the process of socialization, it can be argued that intuitive thinking will cause (or at least be associated with) an increase in belief in God. Three groups working independently from one another tested the above hypothesis and provided evidence that, as belief in God increased, the tendency of reflective thinking decreased (Gervais & Norenzayan, 2012; Pennycook, Cheyne, Seli, Koehler & Fugelsang, 2012; Shenhav, Rand & Greene, 2012). Although more recent findings have suggested that the relationship varies according to different

types of religiosity (Bahçekapili & Yilmaz, 2017) and culture (Gervais et al., 2018), a meta-analysis found a significant, yet weak, negative relationship between the reflective thinking tendency and religious belief (Pennycook, Ross, Koehler & Fugelsang, 2016).

However, findings questioning the causal influence of reflective thinking on belief in God have also emerged (e.g., Farias et al., 2017). While Gervais and Norenzayan (2012) showed in 4 different experiments that activating reflective thinking caused a significant decrease in belief in God, one of their experiments (Study 2) could not be replicated in a study with high statistical power (Sanchez et al. 2017). On the other hand, one of their other experiments (Study 4) was conceptually replicated in a small Turkish sample (Yilmaz, Karadöller & Sofuoğlu, 2016). Yet another study indirectly replicating Gervais and Norenzayan (2012) with a different method observed that reflective thinking caused an increase in intrinsic religiosity in two different American samples (Yonker, Edman, Cresswell & Barrett, 2016). However, none of these studies provided additional evidence that the methods employed to prime reflective thinking did indeed activate reflective thinking. For instance, the first of the three different manipulations used by Gervais and Norenzayan (2012; visual priming), had no effect in the studies of Sanchez et al. (2017) and Deppe et al. (2015). While Yilmaz et al.'s (2016) study revealed an effect for scrambled sentence task, replicating the findings of Gervais and Norenzayan (2012), no effect was detected in another study on moral sensitivity (Yilmaz & Bahçekapili, 2016). Likewise, the cognitive disfluency paradigm (Study 5) used by Gervais and Norenzayan (2012) could not be replicated in a

---

*Department of Psychology, Boğaziçi University, Bebek, 34342 Istanbul, Turkey. E-mail: adil.saribay@boun.edu.tr. ORCID: 0000-0001-7070-7106.

†Department of Psychology, Kadir Has University, Istanbul, Turkey. ORCID: 0000-0002-6094-7162.

‡Department of Psychology, Ludwig-Maximilians-Universität München, Germany. ORCID: 0000-0002-3216-6844.

high-powered study (Meyer et al., 2015) and had no effect in another study with Turkish university students (Yilmaz & Saribay, 2016). The method of manipulation deployed by Yonker et al. (2016) is problematic in itself because it was assumed that getting the participants to work on the analytic conjunction task (Johnson-Laird & Wason, 1970) — normally used to measure reflective thinking tendency — or Stroop test at the beginning of the experiment would increase reflective thinking. In short, there seems to be a problem of consistently priming reflective thinking in the literature. As a result, it is not clear whether these manipulations work as intended, and consequently whether there is truly a causal relationship between reflective thinking and belief in God.

Another study that examined the relationship between reflective thinking tendency and belief in God experimentally was the study conducted by Shenhav et al. (2012), where participants were assigned to either reflective or intuitive thinking conditions using the thought prime technique. Participants in the intuitive thinking condition showed stronger belief in God. However, since the data for this study were collected from mTurk, and since the study has never been replicated, a need for replication in a non-Western culture seems necessary, also taking into consideration the mixed results in the literature mentioned above. Gervais et al. (2018) also mentioned the possibility that the analytic thought-atheism relation holds mostly for religious cultures, where countering societal religious norms requires analytic thought. Hence, in this study, we attempt to replicate this particular study by Shenhav et al. (2012; Study 3) in Turkey, a predominantly Muslim country which has experienced increasing public expression of religiosity in the past decade.

2   Method

This replication was registered at Open Science Framework (OSF). The registration is available at https://osf.io/6cknr/. Data for the present study, a codebook explaining the variables, the analysis script, and the psytoolkit file for the online experiment are provided in the same OSF page.

2.1   Participants and Sample Size Estimation

In the original study by Shenhav et al. (2012; Study 3), 373 mTurk workers were randomly assigned to four conditions. As explained below, we omitted two of these conditions. For sample size estimation, we assumed a small effect size ($d = .20$; Cohen, 1988), set alpha at .05 (one-tailed, see Analysis Plan section below) and power at .90. Using G*Power software (Faul, Erdfelner, Buchner & Lang, 2009), we computed the required total sample of participants to be at least 858 to detect a difference between two conditions in an independent-samples t-test. Our stopping rule was to reach this number after the exclusions stated below. In some

recent analyses, Rand (2018a) observed lower effect sizes; specifically, Cohen's $d = .15$ if analyzed via Tobit regression (see Analysis Plan section below); and Cohen's $d = .06$ if analyzed via t-test. We decided against taking this latter effect size estimate as basis for our study because it yields an infeasible required sample of 11,678 participants with the same parameters as above. Undergraduate students studying at Boğaziçi University (Istanbul) took part in the experiment in exchange for extra course credit. Demographics for the participants who passed the attention check are provided in the *Descriptives* section, below.

2.2   Materials and Procedure

Shenhav et al. (2012; Study 3) used a 2*2 between-subjects design, crossing mindset with outcome valence. In the intuition-negative condition, participants were asked to "write a paragraph (approximately 8–10 sentences) describing a time your *intuition/first instinct* led you in the *wrong* direction and resulted in a *bad* outcome." In the intuition-positive condition section, participants were told to write the same kind of paragraph to "describe a time your *intuition/first instinct* led you in the *right* direction and resulted in a *good* outcome." Likewise, in the reflection-negative condition, participants were asked to "write a paragraph (approximately 8–10 sentences) describing a time *carefully reasoning* through a situation led you in the *wrong* direction and resulted in a *bad* outcome." In the reflection-positive condition, participants were asked to write the same kind of paragraph "describing a time *carefully reasoning* through a situation led you in the *right* direction and resulted in a *good* outcome." It was assumed that reminding participants of a positive (vs. negative) outcome associated with reflective thinking would temporarily reinforce the tendency to use reflection; whereas reminding them of a positive (vs. negative) outcome associated with intuitive thinking would temporarily reinforce the tendency to use intuition. Even though this method was used successfully in many different studies afterwards, a meta-analysis study conducted by Rand (2018a) indicated that only intuition-positive and reflection-positive manipulations produced an effect. Therefore, we chose to use these two conditions that we refer to from hereon as pro-intuition and pro-reflection.

In line with the original experiment, participation was online. The study was presented using PsyToolkit (Stoet, 2010, 2017). Participants were recruited via an e-mail invitation. Upon consenting to participate, they were randomly assigned to pro-intuition or pro-reflection conditions. Participants in the pro-intuition condition received the intuition-positive instructions given above; participants in the pro-reflection condition received the reflection-positive instructions given above. All participants were exposed to two different belief in God measures. The first measure asked "To what extent do you believe in God's existence?" Responses were col-

lected on a Likert-type scale ranging from 1 (not at all) to 10 (certainly). Although the original experiment included a second item ("I have had an experience, which convinced me in God's existence"), we excluded it since it measures recall of an experience rather than belief per se. Instead, we added a more sensitive multi-item religious belief scale. The Intuitive Religious Belief Scale (IRBS) was developed by Gervais and Norenzayan (2012), and adapted to Turkish by Yilmaz and Bahçekapili (2015). This scale consists of five items (e.g., "I believe in God") with a 5-point (1 = strongly disagree, 5 = strongly agree) response scale. Participants were presented with the Faith in Intuition scale (FII) developed by Pacini and Epstein (1999) and adapted into Turkish by Türk and Artar (2014), as a manipulation check (i.e., pro-intuition participants should report greater FII than pro-reflection participants). Items appeared in individually randomized order for these two scales (with the order of scales fixed — IRBS always appeared first). Three IRBS items and three FII items were reverse-coded. Mean of the IRBS and FII items were used in the analyses. Both scales had a Cronbach's α of .9.

Participants were then presented with three Cognitive Reflection Test items (CRT; Frederick, 2005) as a performance-based manipulation check (i.e., pro-reflection participants should perform better on the CRT than pro-intuition participants). CRT is one of the most widely used thinking style measure in the literature (sample item: "A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost? ___cents"). In this question, the correct answer is 5 cents, but there is also an intuitive, low effort answer to this question (10 cents). Responses were scored as correct or incorrect and the total number of correct responses was used in analyses involving CRT.[1] Cronbach's α for the three CRT items was .59, which is typical (Baron, Scott, Fincher & Metz, 2015). These latter two measures (FII and CRT) were not originally used by Shenhav et al. (2012) but we chose to use them for manipulation check purposes (i.e., total CRT score should be higher in the pro-reflection condition and the mean FII score should be higher in the pro-intuition condition).

Next, also unlike Shenhav et al.'s (2012) procedure, as an attention check, participants were asked to demonstrate recognition of the instructions they were given at the beginning of the experiment (i.e., what kind of mental activity they were asked to think about) by choosing from the following options, presented in random order: (a) "carefully reasoning," (b) "using intuition/first instinct," and (c) "I do

---

[1] 40 participants out of the 858 who passed the attention check answered the first CRT question (see Appendix) with "0.05." We had requested the response in "kuruş" (cents) as the unit. This response is senseless in units of kuruş whereas it is actually the correct response in units of lira (the larger unit of Turkish currency akin to dollar). Thus, there is very strong reason to believe that these participants ignored the unit in the question stem and provided their responses in terms of "lira." We considered these responses as correct.

not remember." The first option is correct for those in the pro-reflection condition and the second option is correct for those in the pro-intuition condition. Participants who choose option (c) and those who answer incorrectly were excluded from analyses. Finally, basic demographic information (e.g., age, gender, religious affiliation) was collected. A full version of the experiment design can be found in the appendix. The appendix is in English whereas the actual experiment used direct translations of this material into Turkish (available in the psytoolkit file in OSF).

3 Results

All operations on the data reported below were performed with R, version 3.6.0 (R Core Team, 2019).

3.1 Data Preparation and Cleaning

Participants who did not complete the whole experiment and those who failed to pass the attention check (i.e., those who indicated that they do not remember the instructions at the beginning of the experiment and those who remember the instructions incorrectly; see Materials and Procedures) were excluded from analyses. Additionally, in the original study, Shenhav et al. (2012) excluded from the analysis those who wrote less than 8 sentences in the manipulation phase. However, this can lead to random selection problems as acknowledged by one of the original authors (D. G. Rand, personal communication, May 15, 2018). Therefore, we did not follow this practice.

To reach our desired sample size, we collected data from 1046 participants. 188 of these (18%) failed the check question by either reporting that they do not remember the type of memory that was requested of them in the manipulation phase (36 participants, 3.4% of the total sample) or remembering the wrong type of manipulation (152 participants, 14.5% of the total sample). 858 participants (82%) passed the check and were included in the analyses below. There were 431 and 427 participants in the pro-reflection and pro-intuition conditions, respectively.

3.2 Descriptives

Of these 858 participants, 499 (58.16%) were female and 359 (41.84%) were male. Mean age was 20.97 (SD = 2.87). The majority identified ethnically as Turkish (n = 746, 86.95%). The most common religious affiliation was Muslim (n = 419, 48.83%), followed by agnostic (n = 137, 15.97%), belief in God without organized religion (n = 128, 14.92%), and atheist (n = 114, 13.29%). A small number identified as Christian (n = 1, 0.12%), Buddhist (n = 2, 0.23%) or "other" (n = 57, 6.64%). Table 1 shows the means, standard deviations, and Pearson's correlation coefficients for the main measures, separately in each condition of the experiment.

TABLE 1: Means, standard deviations, and correlations with confidence intervals.

Variable	Pro-Intuition					Pro-Reflection				
	M	*SD*	1	2	3	*M*	*SD*	1	2	3
1. Belief in God	6.68	3.43				6.45	3.46			
2. IRBS mean	3.43	1.31	.90 [.88, .92]			3.35	1.30	.89 [.87, .91]		
3. FII mean	3.35	0.82	.14 [.05, .24]	.14 [.05, .23]		3.22	0.76	.14 [.05, .23]	.14 [.05, .23]	
4. CRT sum	2.24	0.97	−0.05 [−.15, .04]	−0.06 [−.16, .03]	−.16 [−.25, -.06]	2.32	0.92	−.16 [−.25, -.07]	−.15 [−.24, -.05]	−0.06 [−.15, .04]

Note. Values in square brackets indicate the 95% confidence interval for each correlation. IRBS = Intuitive Religious Belief Scale; FII = Faith in Intuition; CRT = Cognitive Reflection Test. See also Footnote 4.

3.3 Confirmatory Analyses

To examine whether there are any significant differences between the two conditions on the main dependent variables (hypotheses 1a and 1b), we conducted Welch's independent samples t-tests (see Delacre, Lakens & Leys, 2017) on both belief measures. Similarly, to check for the effectiveness of the manipulation (hypotheses 2a and 2b), we applied Welch's independent samples t-test on the FII scale and CRT-correct score. These confirmatory tests are directional and were pre-registered. Thus, these t-tests are one-tailed.

Rand (2018a) observed that the majority of his participants used either end of the belief in God scale; reasoned that a t-test in this case may lead to an inaccurate estimate of the true effect size; and used Tobit regression to overcome this potential problem. Tobit regression is suitable for censored data, that is, when observations that are at either end of a scale have unknown higher/lower values due to artificial limits on the scale. To provide results that are comparable to those of Rand (2018a), we also conducted Tobit regression analyses (with condition as predictor) for outcome variables that exhibited clustering at the lower and/or upper ends of the response scale.

3.3.1 Hypothesis 1a: Lower average belief in God score in Pro-Reflection versus Pro-Intuition

Hypothesis 1a stated that participants in the pro-reflection condition would show lower belief in God's existence (as measured by the single item) on average than those in the pro-intuition condition.

Figure 1 (panel A) shows boxplots and data points of belief in God responses across the conditions. A Welch's independent-samples t-test (one-tailed) showed no significant difference between the mean self-reported belief in God in the pro-reflection (*M* = 6.448, *SD* = 3.46) condition and that in the pro-intuition condition (*M* = 6.677, *SD* = 3.43),

$t(856) = −0.973$, $p = 0.166$; 95% CI = $−\infty$ to 0.159, $d = 0.066$.

Belief in God responses were heavily clustered at the upper end of the scale (see Figure 2). Still, testing the same hypothesis with Tobit regression did not change the conclusion of the t-test. The coefficient for condition indicated that the predicted belief in God score was 0.543 (*SE* = 0.43; 95% CIs: −0.3 to 1.388) units higher in the pro-intuition (vs. pro-reflection) condition, but this effect was not significant, $p = 0.207$.

3.3.2 Hypothesis 1b: Lower average IRBS score in Pro-Reflection versus Pro-Intuition

Hypothesis 1b stated that participants in the pro-reflection condition would have a lower average IRBS score than those in the pro-intuition condition. A Welch's independent-samples t-test (one-tailed) showed no significant difference between the mean IRBS score in the pro-reflection condition (*M* = 3.35, *SD* = 1.3) and that in the pro-intuition condition (*M* = 3.426, *SD* = 1.31), $t(855.74) = −0.774$, $p = 0.199$; 95% CI = $−\infty$ to 0.071, $d = 0.058$.

IRBS scores tended to cluster in the upper end of the scale and thus, we tested the same hypothesis with Tobit regression, as well. The above conclusion did not change. The coefficient for condition indicated that the predicted IRBS score was 0.068 (*SE* = 0.106; 95% CIs: -0.139 to 0.275) units higher in the pro-intuition (vs. pro-reflection) condition, but this effect was not significant, $p = 0.522$.

3.3.3 Hypothesis 2a: Lower average FII score in Pro-Reflection versus Pro-Intuition

Hypothesis 2a stated that participants in the pro-reflection condition would have a lower average FII score than those in the pro-intuition condition. Unlike belief in God and IRBS
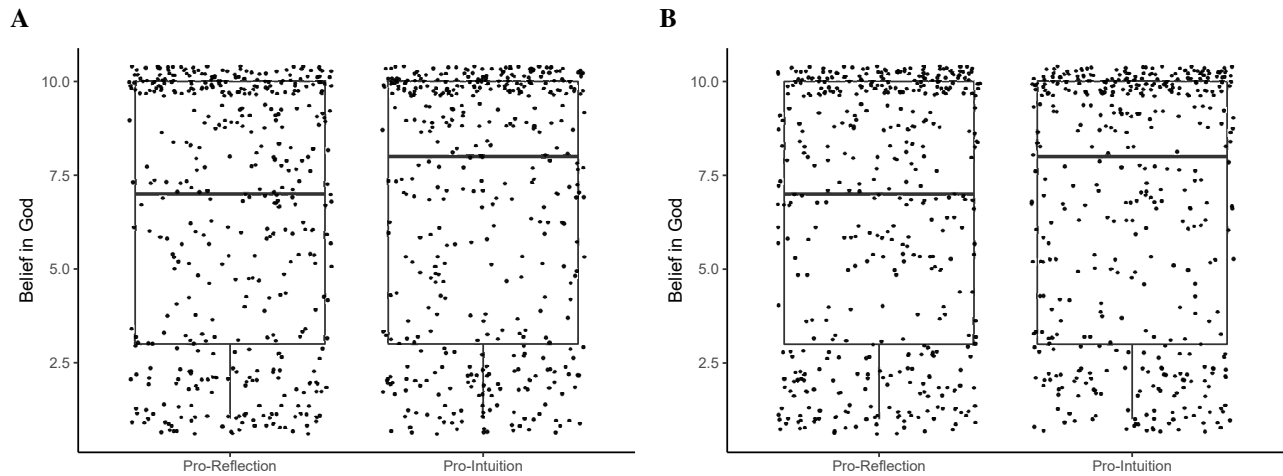
FIGURE 1: Boxplots and data points for Belief in God responses (A) and IRBS mean scores (B).
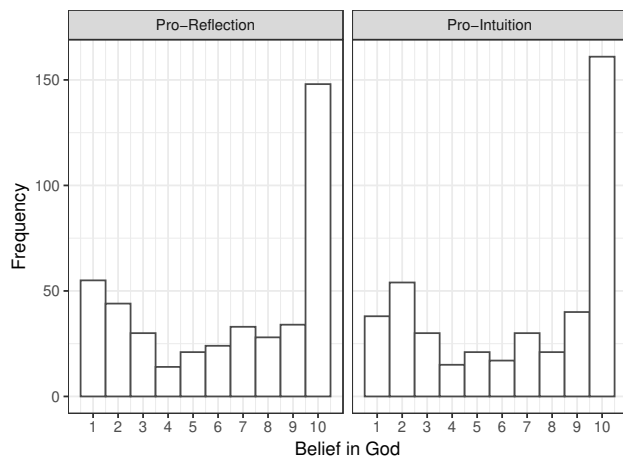

FIGURE 2: Histogram of Belief in God responses in the two conditions of the experiment.

scores, FII scores were normally distributed. A Welch's independent-samples t-test (one-tailed) yielded evidence that the mean self-reported FII score in the pro-reflection condition ($M = 3.222$, $SD = 0.757$) was indeed lower than that in the pro-intuition condition ($M = 3.354$, $SD = 0.822$), $t(848.92) = -2.442$, $p = 0.007$; 95% CI $= -\infty$ to -0.043, $d = 0.167$.

3.3.4 Hypothesis 2b: Higher CRT total score in Pro-Reflection versus Pro-Intuition

Hypothesis 2b stated that participants in the pro-reflection condition would have a higher total CRT score than those in the pro-intuition condition. A Welch's independent-samples t-test (one-tailed) showed no significant difference between the total CRT score in the pro-reflection condition ($M = 2.323$, $SD = 0.923$) and that in the pro-intuition condition (M

$= 2.237$, $SD = 0.968$), $t(853.23) = 1.331$, $p = 0.908$; 95% CI $= -\infty$ to 0.192, $d = 0.091$.

3.4 Exploratory Analyses

In order to test whether the mentioned effect exists only among those who already have an existing belief in God, we repeated the analyses excluding self-reported atheists and agnostics. We also repeated the analyses for only self-reported Muslim participants.[2] Additionally, we planned to examine whether the effect is stronger in those with an already high or low FII and CRT in an exploratory manner, in the event that these scores are not affected by the manipulation. Only CRT fulfilled this criterion (see tests of H2a and H2b, above). Consequently, we conducted a moderated regression analysis to examine whether CRT scores moderated the effect of our IV on the two belief in God DVs.

3.4.1 Analyses Excluding Nonbelievers

We repeated the analyses above excluding nonbelievers (i.e., atheists and agnostics). In the subset of participants who passed the attention check, there were 57 who responded with "other" to the item asking their religious affiliation. Open-ended responses (e.g., "apatheist," "I do not dwell on this matter," "I'm undecided, I cannot answer with certainty," "deist (sometimes atheist)," "I am not in a position to understand whether there is a creator or not") indicated that it would be difficult to categorize many of these "other" responses as either believer or non-believer. Thus, to err on the side of caution, we excluded all participants who responded with "other," in addition to those choosing "atheist"

---

[2]Since we found no reason to change our expectation regarding the direction of effects in the analyses mentioned so far, we kept these tests one-tailed like their confirmatory counterparts.

or "agnostic."[3] This resulted in a sample of 550 participants of which 272 and 278 were in the pro-reflection and pro-intuition groups, respectively. The power of this sample to detect an effect the size of $d = 0.2$ (with $\alpha = .05$) in a one-tailed independent-samples t-test was 0.757. The results below should be interpreted keeping in mind the drop in power.

A one-tailed Welch's independent-samples t-test showed no significant difference between the mean self-reported belief in God in the pro-reflection ($M = 8.592$, $SD = 1.96$) condition and that in the pro-intuition condition ($M = 8.809$, $SD = 1.86$), $t(544.85) = -1.332$, $p = 0.092$; 95% CI = $-\infty$ to -0.051, $d = 0.114$.

Because these participants were all believers, their belief in God scores tended to cluster heavily on the upper end. Still, testing the same hypothesis with Tobit regression did not change this conclusion. The coefficient for condition indicated that the predicted belief in God score was 0.465 ($SE = 0.353$; 95% CIs: -0.228 to 1.157) units higher in the intuition (vs. reflection) condition, but this effect was not significant, $p = 0.188$.

A series of one-tailed Welch's independent-samples t-tests was conducted to test the remaining hypotheses (H1b, H2a, and H2b) on the subsample of believers. There was no significant difference between the mean self-reported IRBS score in the pro-reflection ($M = 4.088$, $SD = 0.882$) condition and that in the pro-intuition condition ($M = 4.186$, $SD = 0.801$), $t(540.47) = -1.365$, $p = 0.086$; 95% CI = $-\infty$ to 0.02, $d = 0.117$. There was, however, evidence to suggest that the mean self-reported FII score in the pro-reflection ($M = 3.319$, $SD = 0.713$) condition was lower than that in the pro-intuition condition ($M = 3.431$, $SD = 0.795$), $t(543.93) = -1.737$, $p = 0.041$; 95% CI = $-\infty$ to -0.006, $d = 0.148$. Finally, there was no significant difference in CRT total scores between pro-reflection ($M = 2.191$, $SD = 0.976$) and pro-intuition ($M = 2.173$, $SD = 0.979$) conditions, $t(547.82) = 0.222$, $p = 0.588$; 95% CI = $-\infty$ to 0.156, $d = 0.019$.

3.4.2 Analyses Including Only Muslims

Gervais et al. (2018) speculated that the relationship between religious disbelief and reflective thinking may be weaker in cultures where religious institutions enjoy weaker influence or in which atheism is institutionalized. This implies that in cultures with stronger (vs. weaker) religious influence, reflective thinking should predict religious disbelief more strongly. At the individual level, this may be taken to suggest that the subsample of participants who identify as Muslim (who presumably were embedded more in subcultures with stronger religious institutions) may exhibit a stronger relation

---

[3]For purposes of comparison and completeness, we replicated the analyses in this section including "other" responders into the subsample of "believers," resulting in 607 cases. This did not change the conclusions except that the effect of condition on the mean FII score was no longer significant, $p = 0.054$.

between religious disbelief and reflectiveness. Thus, confirmatory analyses were also repeated with participants who identified as Muslim ($n = 419$). There were 203 participants in the pro-reflection condition and 216 in the pro-intuition condition. The power to detect an effect the size of $d = 0.2$ (with $\alpha = .05$) in a one-tailed independent-samples t-test was 0.655. Thus, the results below should be interpreted with caution.

Hypothesis 1a was tested with a t-test and Tobit regression, as above. A one-tailed Welch's independent-samples t-test showed no significant difference between belief in God scores in the pro-reflection ($M = 9.291$, $SD = 1.32$) versus pro-intuition ($M = 9.449$, $SD = 1.03$) conditions, $t(382.46) = -1.363$, $p = 0.087$; 95% CI = $-\infty$ to 0.033, $d = 0.134$.

Testing the same hypothesis with Tobit regression yielded the same conclusion. The coefficient for condition indicated that the predicted belief in God score was 0.325 ($SE = 0.36$; 95% CI: -0.38 to 1.03) units higher in the intuition (vs. reflection) condition, but this effect was not significant, $p = 0.366$.

A series of one-tailed Welch's independent-samples t-tests were conducted to test the remaining hypotheses (H1b, H2a, and H2b) on the subsample of Muslims. There was no significant difference between the mean IRBS score in the pro-reflection ($M = 4.461$, $SD = 0.583$) condition and that in the pro-intuition condition ($M = 4.482$, $SD = 0.543$), $t(409.72) = -0.387$, $p = 0.35$; 95% CI = $-\infty$ to 0.07, $d = 0.038$. There was also no significant difference between the mean FII score in the pro-reflection condition ($M = 3.286$, $SD = 0.7$) and that in the pro-intuition condition ($M = 3.389$, $SD = 0.774$), $t(416.4) = -1.433$, $p = 0.076$; 95% CI = $-\infty$ to 0.016, $d = 0.14$. Finally, there was no significant difference in CRT total scores between pro-reflection ($M = 2.246$, $SD = 0.954$) versus pro-intuition ($M = 2.208$, $SD = 0.964$) conditions, $t(415.88) = 0.405$, $p = 0.657$; 95% CI = -∞ to 0.192, $d = 0.04$.

3.4.3 Moderation by CRT scores

We examined whether the effect is stronger for participants with an already high or low analytical thinking tendency. To examine whether CRT performance moderated the effect of our manipulation on belief in God, a regression analysis was conducted including condition (with the pro-reflection condition serving as the reference category), CRT total score, and their interaction as predictors, on the whole sample of participants included in the confirmatory analyses, $F(3, 854) = 4.461$, $p = 0.004$, $Adj. R^2 = 0.012$. There was no interaction between CRT and condition, $\beta = 0.115$, $SE = 0.068$, $p = 0.093$.

We examined the same moderation using IRBS scores as the outcome variable because IRBS is conceptually similar to belief in God. This regression model yielded similar results as above, $F(3, 854) = 3.924$, $p = 0.008$, $Adj. R^2 = $

0.01. There was no interaction between CRT and condition, $\beta = 0.089$, $SE = 0.068$, $p = 0.192$.[4]

4   Discussion

This replication attempt was, to our knowledge, the first high-powered experiment investigating the roles of intuition and reflection on religious belief in a non-WEIRD sample (Henrich, Heine, & Norenzayan, 2010) from a predominantly Muslim country with a secular tradition but where religious institutions currently reign supreme (Çarkoğlu & Kalaycıoğlu, 2009). The findings show that the experimental manipulation used here to activate either intuitive or reflective thinking had a statistically significant effect on neither the single item belief in God question nor a reliable religiosity scale, although one of the manipulation checks (FII, but not CRT) showed evidence of its effectiveness. Similarly, no statistically significant effect of the manipulation was found among believers (i.e., when excluding non-believers). These results do not support the *intuitive belief hypothesis* that intuition increases and that reflection decreases religious belief (Gervais & Norenzayan, 2012; Shenhav et al., 2012; Yilmaz et al., 2016), and are in line with recent failed replication attempts (Farias et al., 2017; Sanchez, Sundermeier, Gray, & Calin-Jageman, 2017; Yonker et al., 2016). However, it should be noted that the pattern of findings were in the predicted direction.

A question arises here: Is this a replication or a manipulation failure? The results showed that the manipulation influenced participants' differential reliance on intuition versus reflection on a self-report measure (FII), but not on a performance-based measure (CRT). This discrepancy between self-reported and actual performance might point to a demand effect problem. That is, especially given the explicit and transparent nature of the manipulation, it is possible that participants strategically responded to subsequent measures in line with their naïve theory of how anyone in that particular mindset (intuition or reflection) would respond; instead of the manipulation actually activating an authentic mindset with real (non-strategic) downstream consequences on beliefs. In fact, a brief examination of open-ended debriefing responses reveals that many, if not most, participants iden-

tified the study as being about the "religious belief and the tension between logic and intuition." However, the data do not allow us to examine the presence, degree, and direction of strategic responses that may have accompanied such insight into the nature of the experiment. It is also possible that the effect of the manipulation dissipated by the time the CRT items were presented (always after FII), which would explain the discrepancy. Future studies should use other techniques that are known to reliably influence actual performance to activate reflection, such as time-delay with prompts to think reflectively (see Yilmaz & Isler, 2019).

In contrast to the null result indicating no effect of manipulation on the main dependent variables, there was a statistically significant negative relation between the tendency to think reflectively (as evidenced by the scores on CRT) and religious belief (both measures) in one of the two conditions, in line with the previous literature (Bahcekapili & Yilmaz, 2017; Gervais & Norenzayan, 2012; Gervais et al., 2018; Pennycook et al., 2012; Saribay & Yilmaz, 2017; Stagnaro, Ross, Pennycook & Rand, 2019; Yilmaz & Saribay, 2016). Interestingly, the same discrepancy between experimental and correlational findings was observed in another recent study (Yilmaz & Isler, 2019). As those authors also speculated, it is possible that these methods capture different psychological mechanisms — placing one's established beliefs under scrutiny versus long-term attraction to subcultures and beliefs more consistent with one's cognitive inclinations, respectively. Future studies should investigate the reasons for this discrepancy more directly.

What could be the cause of the contradictory findings in the literature? First, it is possible that the intuitive religious belief hypothesis is invalid. For instance, contrary to this hypothesis, a recent high-powered study demonstrated that reflection increases belief in God among non-believers through self-questioning (Yilmaz & Isler, 2019). Second, as mentioned before, the statistically non-significant effects were in the direction predicted by the intuitive religious belief hypothesis. Therefore, it is possible that there is a true effect in the hypothesized direction but it is too small to be captured even with a large sample size such as ours. In their meta-analysis, Pennycook et al. (2016) found a significant negative correlation between analytic thinking and religious belief — albeit with a very small effect size ($r = .15$). Capturing such small effects experimentally would require even stronger manipulation techniques (e.g., comprehensive analytic thought training sessions) or much larger sample sizes. In fact, Rand (2018b) recently argued that stronger manipulations may be required to observe the effects of intuitive thought, although his focus was on replicating these kinds of effects in the increasingly experienced sample of mTurkers.

A third possibility is the existence of boundary conditions. In particular, we conjecture that the belief in God question more likely probes stable opinions which are difficult to change with an experimental manipulation since they

---

[4]One may wish to know about the main effect of CRT in these analyses. Since regression estimates change depending on which condition is taken as the reference category, it is easier to appreciate the relation between CRT total score and the outcome variables used here (belief in God and IRBS mean score) by examining the correlations in Table 1 instead. Even though those correlations appear to be stronger in the Pro-Reflection condition, consistent with the lack of interaction effects reported in these regression analyses, they do not differ when compared with Fisher's r-to-z transformation with the "cocor" package in R (Diedenhofen & Musch, 2015). For the CRT-belief in God correlations: $z = 1.33$, $p = 0.184$. Since this pair of correlation coefficients is the one that is the farthest apart in Table 1, this implies that no other pair of correlation coefficients (keeping both variables the same) differ across the two conditions.

reside in memory as defining characteristics of one's personal identity over the course of one's life. Consistent with this interpretation, 46.85% of all participants in our study chose either the minimum or the maximum possible response on this question. The type of cognitive process manipulations used in the literature may not be strong enough to influence stable opinions. To date, only one high-powered experiment (Yilmaz & Isler, 2019) found an effect on such stable opinions using a stronger manipulation (time-limit manipulation with prompts in a within-subjects design), but in the *opposite* direction (i.e., reflection increases belief in God) as that put forth by the intuitive belief hypothesis. It is possible that strong cognitive style manipulations have more influence on contextualized opinions. This implies that, instead of asking about belief in God, it may be more sensible to ask about the involvement of a supernatural agent on a mysterious event such as disappearance of Airlines Flight MH370 (see also a similar distinction between stable and contextualized opinions, Talhelm, 2018; Talhelm et al., 2015; Yilmaz & Saribay, 2017).

Another possible moderating variable is the individual differences in baseline levels of reflective thinking tendency. It may be easier to observe the effects of cognitive process manipulations using a pretest-posttest design, where the reflective thinking tendency of participants is measured prior to experimental manipulation. Our exploratory analyses examined whether our intuition manipulation may have a stronger effect on people who naturally tend to rely on reflection relative to those who are already predisposed to making intuitive decisions. This did not appear to be the case. However, we measured reliance on intuition and reflection after the experimental manipulation. Future studies should measure reflective thinking tendency prior to the experimental manipulation, to obtain baseline levels not contaminated by the manipulation.

In summary, this and similar experiments that apparently failed to support the intuitive religious belief hypothesis suggest that the previously found effects are not easily replicable, even with high-powered studies. As in the correlational investigation of the hypothesis (i.e., Gervais et al., 2018), multi-lab experiments are needed to shed further light on the relationship between religious belief and cognitive style.

References

Bahçekapili, H. G., & Yilmaz, O. (2017). The relation between different types of religiosity and analytic cognitive style. *Personality and Individual Differences*, *117*, 267–272.

Baron, J., Scott, S., Fincher, K., & Metz, S. E. (2015). Why does the cognitive reflection test (sometimes) predict utilitarian moral judgment (and other things)? *Journal of*

Applied Research in Memory and Cognition, 4(3), 265–284.

Çarkoğlu, A., & Kalaycıoğlu, E. (2009). *The rising tide of conservatism in Turkey*. New York, NY: Palgrave Macmillan US.

Delacre, M., Lakens, D., & Leys, C. (2017). Why psychologists should by default use Welch's *t*-test instead of student's *t*-test. *International Review of Social Psychology, 30*(1), 92–101.

Deppe, K. D., Gonzalez, F. J., Neiman, J. L., Jacobs, C., Pahlke, J., Smith, K. B., & Hibbing, J. R. (2015). Reflective liberals and intuitive conservatives: A look at the Cognitive Reflection Test and ideology. *Judgment and Decision Making*, *10*(4), 314–331.

Diedenhofen, B. & Musch, J. (2015). cocor: A comprehensive solution for the statistical comparison of correlations. *PLoS ONE, 10*(4): e0121945.

Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, *8*(3), 223–241.

Farias, M., Mulukom, V., Kahane, G., Kreplin, U., Joyce, A., Soares, P., ... & Möttönen, R. (2017). Supernatural belief is not modulated by intuitive thinking style or cognitive inhibition. *Scientific Reports*, *7*, 15100.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic perspectives*, *19*(4), 25–42.

Gervais, W. M., & Norenzayan, A. (2012). Analytic thinking promotes religious disbelief. *Science*, *336*(6080), 493–496.

Gervais, W. M., van Elk, M., Xygalatas, D., Mckay, R. T., Aveyard, M., Buchtel, E. T., ... & Svedholm-Häkkinen, A. M. (2018). Analytic atheism: A cross-culturally weak and fickle phenomenon? *Judgment and Decision Making, 13*(3), 268–274.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*(2–3), 61–83.

Johnson-Laird, P. N., & Wason, P. C. (1970). A theoretical analysis of insight into a reasoning task. *Cognitive Psychology*, *1*(2), 134–148.

Meyer, A., Frederick, S., Burnham, T. C., Guevara Pinto, J. D., Boyer, T. W., Ball, L. J., ... & Schuldt, J. P. (2015). Disfluent fonts don't help people solve math problems. *Journal of Experimental Psychology: General*, *144*(2), e16-e30.

Pacini, R., & Epstein, S. (1999). The relation of rational and experiential information processing styles to personality, basic beliefs, and the ratio-bias phenomenon. *Journal of Personality and Social Psychology*, *76*(6), 972–987.

Pennycook, G., Cheyne, J. A., Seli, P., Koehler, D. J., & Fugelsang, J. A. (2012). Analytic cognitive style predicts religious and paranormal belief. *Cognition*, *123*(3), 335–346.

Pennycook, G., Ross, R. M., Koehler, D. J., & Fugelsang, J. A. (2016). Atheists and agnostics are more reflective than religious believers: Four empirical studies and a meta-analysis. *PloS One*, *11*(4), e0153039.

R Core Team (2019). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

Rand, D. G. (March, 2018a). The relationship between belief in God, intuition, and prosociality. *Talk given at the 19th Annual Convention of the Society for Personality and Social Psychology: Psychology of Religion & Spirituality Preconference, Atlanta, USA.*

Rand, D. G. (2018b). Non-naïvety may reduce the effect of intuition manipulations. *Nature Human Behaviour, 2*, 602.

Sanchez, C., Sundermeier, B., Gray, K., & Calin-Jageman, R. J. (2017). Direct replication of Gervais & Norenzayan (2012): No evidence that analytic thinking decreases religious belief. *PloS One*, *12*(2), e0172636.

Saribay, S. A., & Yilmaz, O. (2017). Analytic cognitive style and cognitive ability differentially predict religiosity and social conservatism. *Personality and Individual Differences, 114*, 24–29.

Shenhav, A., Rand, D. G., & Greene, J. D. (2012). Divine intuition: Cognitive style influences belief in God. *Journal of Experimental Psychology: General, 141*(3), 423–428.

Stagnaro, M. N., Ross, R. M., Pennycook, G., & Rand, D. G. (2019). Cross-cultural support for a link between analytic thinking and disbelief in God: Evidence from India and the United Kingdom. *Judgment and Decision Making, 14*(2), 179–186.

Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods, 42*(4), 1096–1104.

Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology, 44*(1), 24–31.

Talhelm, T., Haidt, J., Oishi, S., Zhang, X., Miao, F. F., & Chen, S. (2015). Liberals think more analytically (more "WEIRD") than conservatives. *Personality and Social Psychology Bulletin, 41*(2), 250–267.

Talhelm, T. (2018). Hong Kong liberals are WEIRD: Analytic thought increases support for liberal policies. *Personality and Social Psychology Bulletin, 44*(5), 717–728.

Türk, E. G., & Artar, M. (2014). Adaptation of the rational experiential inventory: Study of reliability and validity. *Ankara University Journal of Faculty of Educational Sciences*, *47*(1), 1–18.

Yilmaz, O., & Bahçekapili, H. G. (2015). Without God, everything is permitted? The reciprocal influence of religious and meta-ethical beliefs. *Journal of Experimental Social Psychology*, *58*, 95–100.

Yilmaz, O., & Bahçekapili, H. G. (2016). Supernatural and secular monitors promote human cooperation only if they remind of punishment. *Evolution and Human Behavior*, *37*(1), 79–84.

Yilmaz, O., & Isler, O. (2019). Reflection increases belief in God through self-questioning among non-believers. *Judgment and Decision Making, 14*(6), 649–657.

Yilmaz, O., Karadöller, D. Z., & Sofuoglu, G. (2016). Analytic thinking, religion, and prejudice: An experimental test of the dual-process model of mind. The International *Journal for the Psychology of Religion*, *26*(4), 360–369.

Yilmaz, O., & Saribay, S. A. (2016). An attempt to clarify the link between cognitive style and political ideology: A non-western replication and extension. *Judgment and Decision Making*, *11*(3), 287–300.

Yilmaz, O., & Saribay, S. A. (2017). Analytic thought training promotes liberalism on contextualized (but not stable) political opinions. *Social Psychological and Personality Science*, *8*(7), 789–795.

Yonker, J. E., Edman, L. R. O., Cresswell, J., & Barrett, J. L. (2016). Primed analytic thought and religiosity: The importance of individual characteristics. *Psychology of Religion and Spirituality*, *8*(4), 298–308.

Appendix

Experimental Manipulations

"Write a paragraph (approximately 8-10 sentences) describing a time *carefully reasoning* through a situation led you in the *right* direction and resulted in a *good* outcome".

"Write a paragraph (approximately 8-10 sentences) describing a time your *intuition/first instinct* led you in the *right* direction and resulted in a *good* outcome".

Belief in God Question

To what extent do you believe in God's existence?

1	2	3	4	5	6	7	8	9	10
Not at all									Certainly

Intuitive Religious Belief Scale

Response format:

1	2	3	4	5
Absolutely Disagree				Absolutely Agree

1) I believe in God.
2) When I am troubled, I feel the need to seek help from God.
3) People think they talk to God when they are praying but in fact they just talk to themselves. (R)
4) Religion does not make sense to me. (R)
5) Religion plays no role in my daily life. (R)

Faith in Intuition Scale

Response format:

1	2	3	4	5
Absolutely Disagree				Absolutely Agree

1) Using my gut feelings usually works well for me in figuring out problems in my life.
2) Intuition can be a very useful way to solve problems.
3) I hardly ever go wrong when I listen to my deepest gut feelings to find an answer.
4) I believe in trusting my hunches.
5) I often go by my instincts when deciding on a course of action.
6) I like to rely on my intuitive impressions.
7) I don't like situations in which I have to rely on intuition. (R)
8) I don't have a very good sense of intuition. (R)
9) I don't think it is a good idea to rely on one's intuition for important decisions. (R)
10) I trust my initial feelings about people.
11) I tend to use my heart as a guide for my actions.

CRT

1) A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost? (Cents)
2) If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? (Minutes)
3) In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? (Days)

Manipulation Check

At the beginning of the study, we asked you to write a paragraph about a memory of yours in which a particular thinking style led you in the right direction. Which of the below options was that particular thinking style? (If you do not remember clearly, please choose the option "I do not remember").
- ○ Carefully reasoning
- ○ Using intuition/first instinct
- ○ I do not remember

Demographic Questions

Which of the following is written in the "gender" section of your identity card: ○ Female ○ Male

Your age: _____

We often hear about "left" and "right" in political matters. Below is a scale from 1 to 7, 1 being the most "left" and 7 being the most "right". Where would you place your opinions on this scale?

1	2	3	4	5	6	7
Extreme left			Moderate			Extreme right

How would you describe your family's income during your upbringing? ○ Very poor ○ Poor ○ Moderate ○ Good ○ Very good

Your ethnicity: ○ Turkish ○ Kurdish ○ Armenian ○ Greek ○ Arabic ○ Balkans ○ Caucasian ○ Jewish ○ Other

The place you have lived for the longest time is: ○ Metropolis ○ City ○ Big Town ○ Town ○ Village

Please select the option that best describes your religious affiliation: ○ Atheist ○ Agnostic ○ Muslim ○ Belief in God without any organized religion ○ Christian ○ Jewish ○ Buddist ○ Other

Suspiciousness Probes (open ended)

What do you think is being investigated in this study? Did you notice any kind of connection between the different parts of the study?

Did you experience any kind of difficulty during participation (e.g., failure to understand a question or section)? Is there anything you would like the researcher to know?