Replication of

Analytic Thinking Promotes Disbelief

Gervais, W.M. / Norenzayan, A. (2012).
Science, 336, 493-496.

Replication Author(s)
Nick Buttrick, Anup Gampa, Lilian Hummer, & Brian A. Nosek

In a priming experiment with Canadian undergraduates, Gervais and Norenzayan (2012) randomly-assigned participants to one of two conditions before assessing their belief in God. In the Analytic-prime condition, participants viewed 4 pictures of Rodin's *The Thinker*, while participants in the Control condition viewed 4 pictures of Myron's *Discobolus*. Compared to the control condition, participants in the analytic-prime condition reported lower levels of belief in God. The paper included 5 studies. Study 1 is a correlational study, studies 2-4 primed analytic thinking, through images (study 2), or through a scrambled-sentence task (studies 3-4), and study 5 used disfluent fonts to increase analytic thought. Study 2 is the first study in the paper reporting experimental treatment effects, so it was selected for replication.

Hypothesis to replicate and bet on:

Priming analytic thinking via images of "The Thinker" increases religious disbelief compared to viewing control images of a visually similar artwork; a *t*-test, $p<.05$ using a two-tailed test.

Original test statistics: N=57 (31 in Control condition, 26 in Disbelief condition); Control belief in god (100-pt scale): M=61.55, SD=35.68; Disbelief: M=41.42, SD= 31.47; t(55)=2.24; p=0.029 (reported as p = 0.03).

Power Analysis and Criteria for Replication: First Data Collection

The original sample size was 62 observations, with 5 dropped from main analyses due to suspicion of the manipulation (all 5 in the disbelief condition). Final sample size was 57, and the standardized effect size measured as r was 0.289. To have 90% power to detect 75% of the original effect size, a sample size of 224 is required (after exclusion criteria have been met). The criteria for replication is an effect in the same direction as the original study and a p-value<0.05 (two-sided test).

Power Analysis and Criteria for Replication: Second Data Collection

If the original result is not replicated in the first data collection, a second data collection of 290 additional individuals (after inclusion criteria have been met) will be carried out so that the total sample size is 514 (after inclusion criteria have been met). If a second data collection is carried out, it will be tested if the original result replicates in the pooled sample of the first and second data collection.

To have 90% power to detect 50% of the original effect size, a sample size of 514 is required; i.e. a sample size of 290 in the second data collection to have a sample size of 514 in total for the first and second data collection pooled. The criteria for replication is an effect in the same direction as the original and a p-value<0.05 (two-sided test) in the pooled data.

Sample

The sample size in the first data collection will consist of 224 individuals from the Charlottesville, Virginia community. Participants will be recruited on grounds at the University of Virginia or from popular community locations in Charlottesville such as the downtown pedestrian mall. Participants will be compensated $5 directly by the experimenter, or will be recruited using the University of Virginia participant pool for research credit.

If the original result is not replicated in the first data collection (two-sided p-value <0.05 in the original direction), a second data collection of 290 additional individuals from the same population will be carried out so that the total sample size is 514.

Materials

We will run the original Qualtrics script used in the original study, provided to us by the original authors. This script will incorporate the original questions and 8 photographs, as well as the same measure of belief in God [100-point scale, anchored at 0 = God definitely does not exist; and 100 = God definitely exists]. Materials are based on page 4 of the Supplementary Information. The experiment will be in English as in the original study.

Procedure

We will follow the procedure described in the original article. The following summary of the experimental procedure is based on page 494 of the main article, page 4 of the Supplementary Information, and direct feedback provided by the original authors.

Participants will sit in front of a computer that delivers fully-automated instructions. Participants will be randomly assigned by the computer into either the Analytic or Control conditions. In both conditions, participants will be instructed to view 4 slightly different images of a sculpture; in the Analytic condition, they will see 4 images of Rodin's *The Thinker*, while in the Control condition, they will see 4 images of the *Discobolus* of Myron. Instructions will tell participants to spend at least 30 seconds looking at each image before moving on to the rest of the experiment. Once finished with the images, participants will be sent to an ostensibly separate task, where they fill out demographic information including the main DV. The main DV is to rate their belief in God, from 0 (God definitely does not exist) to 100 (God definitely exists). Following a filler task, participants will complete a funneled debrief to check for suspicion that the two parts of the experiment were somehow connected.

Analysis

The analysis will be performed exactly as in the original article. First, anyone expressing suspicion in the debriefing that the two parts of the experiment were really connected will be dropped from the analysis. Questions for the suspicion check in the funneled debriefing have been provided by the original authors. No other exclusion rules were identified in the original

study, so we will include all other participants that respond to the dependent variable. On the remaining sample, condition differences on the belief-in-God DV are analyzed using a two-tailed t-test.

The results will first be estimated based on the first data collection. If the original result is replicated in the first data collection (a two-sided p-value <0.05 in the same direction as the original study), the second data collection will not be carried out. If the original result is not replicated in the first data collection, a second data collection will be carried out. The above statistical test will then be estimated for the pooled sample of the first and second data collection to test if the original result replicated (a two-sided p-value <0.05 in the same direction as the original study).

Differences from Original Study

The replication procedure is identical to that of the original study, with some unavoidable deviations. The replication will be performed in the fall of 2016 or spring of 2017 in Charlottesville, Virginia, whereas the data in the original study was carried out at the University of British Columbia, during the 2009-2010 school year (as indicated by the original authors). As such, as in all replications, the sample, recruiting, and setting are different from the original study. There are no claims in the original article that suggest that these deviations are material for the tested effects. Nevertheless, we have sought review before conducting the replication to confirm.

The original paper contains five studies: for the replication the focus is only on study 2 following the project protocol to select the first study in the paper reporting treatment effects.

Replication Results for the First Data Collection (90% power to detect 75% of the original effect size)

Participants 236 students at the University of Virginia participated in the first data collection for experimental credit. All were recruited from the Department of Psychology's Participant Pool. Sixty percent of the participants were female, with an average age of 18.99 (SD = 1.2). Twelve participants were excluded from analysis by a coder blind to condition due to reporting suspicion that the pictures and their self-reported belief in God were somehow connected, leaving the planned sample of 224 participants. There was no difference in the exclusion rate between conditions: $X^2[1]=0.088$, $p = 0.767$.

The hypothesis was not supported by the data. Participants primed with analytic thought (n=111) did not differ in their reported belief in God as compared to participants primed with a control stimulus (n=113): Analytic M=62.78, SD=36.97; Control M=58.82, SD=35.71; t(223) = 0.815, $p = 0.416$, $r = 0.05$ [95% CI: -0.08, 0.19].

Replication Results for the First and Second Data Collection Pooled (90% power to detect 50% of the original effect size)

Participants: 553 students at the University of Virginia participated in the pooled collection, for experimental credit. All were recruited from the Department of Psychology's Participant Pool. Sixty-two percent of participants were female, with an average age of 18.95 (SD = 1.11). 22 participants were excluded from analysis by a coder blind to condition due to

reporting suspicion that the pictures and their self-reported belief in God were somehow connected, leaving a sample of 531 participants. There was no difference in the exclusion rate between conditions: $X^2[1]=2.29$, $p = 0.130$.

The hypothesis that priming analytic thinking would lead to greater disbelief in God was not supported by the data. Participants primed with analytic thought (n=262) did not differ in their reported belief in God as compared to participants primed with the control stimuli (n=269): Analytic M=62.94, SD=36.40; Control M=60.38, SD=35.87; t(552) = 0.817, $p = 0.414$, $r = 0.04$ [95% CI: -0.05, 0.12].

Since we collected more participants in the second than originally specified, we conducted an additional analysis of just the first 514 participants. Results did not differ meaningfully from the full pooled data. Participants primed with analytic thinking (n=256) did not differ in their reported belief in God as compared to participants primed with the control stimuli (n=258): Analytic M=62.73, SD=36.46; Control M=60.05, SD=36.01; t(513)=0.838, $p = 0.402$, $r = 0.04$ [95% CI: -0.05, 0.12].

We also collected an exploratory analysis of the full pooled data including the participants excluded because of suspicion about the primes. The results were unchanged. Participants primed with analytic thinking (n=277) did not differ in their reported belief in God as compared to participants primed with the control stimuli (n=276): Analytic M=62.53, SD=36.51; Control M=60.51, SD=35.91; t(552)=0.655, $p = 0.513$, $r = 0.03$ [95% CI: -0.06, 0.11].

Unplanned Protocol Deviations

There were no unplanned protocol deviations.

Discussion

Despite high-power and careful adherence to the original study design with feedback from the original authors, we did not observe an effect consistent with the original study. Our effect size was in the same direction as the original but much smaller (r=.04 compared to r=.29), and non-significantly different from zero. The original effect size did not fall within the confidence interval of the replication. This failure to replicate suggests caution about the reliability of the original result, but does not definitively suggest that the original result was a false positive. There could be as yet unidentified differences between the original and replication methodology that are critical for observing this effect.

References

Gervais WM / Norenzayan A. (2012): *Analytic Thinking Promotes Religious Disbelief*, Science, 336, pp. 493-496.