



Atheist primes reduce religiosity and subjective wellbeing

Johannes Haushofer ^{a,b,c} and James Reisinger^{b,d}

^aPrinceton University, Princeton, USA; ^bBusara Center for Behavioral Economics, Nairobi, Kenya; ^cNational Bureau for Economic Research, Cambridge, USA; ^dHarvard University, Cambridge, USA

ABSTRACT

In recent years, atheism has grown in popularity, partly inspired by the rise to prominence of a group of public intellectuals called the “New Atheists” who argue against religion in public fora. What are the social consequences of this development? We test in a laboratory study in Kenya whether exposure to atheist arguments affects self-reported and implicit religiosity, subjective wellbeing, and self-reported tolerance of different social groups. We find a significant negative effect of emotional arguments against religion on both self-reported and implicit measures of religiosity, especially among men, but no effect of scientific appeals. Subjective wellbeing is strongly reduced after emotional atheist messages, again especially among men, suggesting that emotional atheist arguments may have a wellbeing cost. We find no effects of atheism messages on self-reported tolerance. Together, these results suggest that emotional atheist arguments reduce religiosity and subjective wellbeing, especially among men.

ARTICLE HISTORY

Received 15 April 2017

Accepted 6 November 2017


KEYWORDS


Atheism; wellbeing;
laboratory experiment;
priming

1. Introduction

Public intellectuals (e.g. “New Atheists” such as Richard Dawkins, Christopher Hitchens, Sam Harris, and Daniel Dennett) have long and vigorously argued that religion is not only factually inaccurate, but also socially and personally harmful. However, it is unclear whether such arguments can actually influence an individual’s religiosity. In addition, it is unknown whether such advocacy, and atheism in general, has beneficial or detrimental social effects. For instance, diminished religious beliefs might increase or decrease one’s own subjective wellbeing, or one’s tolerance of others. In the present study, we examine whether exposure to arguments against religion can in fact influence an individual’s religiosity, wellbeing, and tolerance of others. We report the results of an experiment conducted among university students in Nairobi, Kenya, in which we expose participants to explicit atheist messages. Participants watch short videos presenting either scientific or emotional arguments against the existence of God. We then measure both self-reported and implicit religiosity, as well as subjective wellbeing and tolerance of others.

Some recent studies have posed similar questions. Concerning the effect of religious or scientific messages on religiosity, a few studies have shed light on this question. Shariff, Cohen, and Norenzayan (2008) found that rational, scientific atheist messages reduced self-reported and implicit religiosity in a Western setting. Preston and Epley (2009) showed that arguments in favor of science as an ultimate explanation of the universe led to more negative evaluations of God (and more positive evaluations of science), while arguments in favor of God as an ultimate explanation had the opposite effect. Relatedly, Preston, Ritter, and Hepler (2013) found that exposure to messages that support

CONTACT Johannes Haushofer  haushofer@princeton.edu

 Supplemental online material for this article can be accessed at <https://doi.org/10.1080/2153599X.2018.1436585>.

© 2018 Informa UK Limited, trading as Taylor & Francis Group

mechanistic explanations of the mind reduced beliefs in the “soul.” Together, previous evidence suggests that scientific messages can reduce religiosity. However, it is also conceivable that a belief-threatening message would induce a reinforcement of the religious belief (the “backfire effect” Nyhan and Reifler (2010)). More broadly, the effect of messages on religious belief speaks to the debate about whether religious beliefs are “factual” and respond to evidence, or whether they are intuitive and respond to authority (Levy, 2017; Van Leeuwen, 2014).

Regarding the effect of religion on subjective wellbeing, cross-sectional evidence suggests that religiosity is positively correlated with subjective wellbeing (for a discussion, see Deaton & Stone, 2013, or Diener, Tay, & Myers, 2011). If these relationships are at least partly reflective of a positive causal effect of religion on these outcomes, one might expect a negative effect of exposure to atheism on these outcomes. A similar hypothesis is suggested by evidence showing that worldview threats can lower self-esteem (Major, Kaiser, O’Brien, & McCoy, 2007) and increase accessibility of death thoughts (Schimmel, Hayes, Williams, & Jahrig, 2007).

Concerning the effect of religious messages on tolerance, both theory and previous evidence suggests that worldview threats can lead to intolerance of others (Bassett & Connelly, 2011; Greenberg, Simon, Pyszczynski, Solomon, & Chatel, 1992). In line with this hypothesis, a number of recent studies have found that religious messages lead to more negative attitudes towards both racial and religious out-groups (Greenberg et al., 1990; Johnson, Rowatt, & LaBouff, 2010; LaBouff, Rowatt, Johnson, & Finkle, 2012; Ramsay, Pang, Shen, & Rowatt, 2014). Relatedly, laboratory studies have found that explicit religious messages and implicit religious primes lead to an increase in pro-social behavior and cooperation towards the in-group (Norenzayan & Shariff, 2008; see Shariff, Willard, Andersen, & Norenzayan, 2015, for a review).

The present study extends these previous findings in several ways. First, we measure exposure both to scientific argumentation and to emotional appeals. The motivation for this distinction is twofold: on one hand, religious belief has been shown to be both intimately entwined with the desire to explain the physical world (Bloom, 2007) and with an emotional desire for solace and comfort in confronting suffering (Ano & Vasconcelles, 2005; Greenberg, Solomon, & Pyszczynski, 1997). On the other, emotional appeals are more effective at changing attitudes and behavior than appeals to reason (Smith & De Houwer, 2015), especially when the target attitudes and behaviors are also emotional (Fabrigar & Petty, 1999). We examine the effect of exposure to anti-religious arguments on religiosity both through self-reported measures and through a Single-Target Implicit Association Test (ST-IAT) measuring the degree to which individuals associate religious concepts with the concepts “real” versus “imaginary.” The motivation for these complementary measures is that if religious belief is not intuitive, but factual, we might not expect effects on a task that taps implicit attitudes such as the ST-IAT. The combination of measures can thus speak to the debate about the intuitive versus factual nature of belief (Levy, 2017; Van Leeuwen, 2014). In addition, the ST-IAT may be less susceptible to demand effects, and more sensitive to subtle belief changes, than self-report measures (Egloff & Schmukle, 2002; Nosek, Greenwald, & Banaji, 2005; Shariff et al., 2008). Third, our target population is university students in Kenya, while most previous studies use Western samples. While our participant pool is not entirely ignorant of atheist arguments, we generally find qualitatively that individuals have had relatively little exposure to, and typically show less acceptance of, these arguments compared to students in other contexts. We might therefore predict that our subject pool is less susceptible to atheist arguments because existing attitudes are more engrained. Finally, we extend previous work by adding outcome measures beyond religiosity, in particular, subjective wellbeing and tolerance of others. The goal of these additions is to begin to ask how exposure to atheist arguments might affect wellbeing and social capital.

2. Study design

2.1. Recruitment and experimental structure

The study was conducted at the Busara Center for Behavioral Economics (Busara) in Nairobi, Kenya, a laboratory facility for social science studies. Busara maintains an active participant pool of more

than 11,000 Nairobi residents. For the present study, 318 participants who had previously signed-up to be part of the Busara subject pool were recruited fusing SMS and phone calls and were informed that they would be paid KES 300 (USD 7.20 PPP)¹ for participation. Participants were told that they were invited to participate in a study about their behavior and preferences.

Recruitment was limited to university students to ensure comprehension of the arguments contained in messages, and to ensure that all participants could participate in the writing task after the messages. In addition, we controlled for heterogeneity in religious background by omitting individuals identified as ethnically Nubian from recruitment. Since most Muslims in the Busara participant pool belong to this ethnicity, we believed this to be the best way to restrict the sample to Christians while avoiding any issues of self-selection that might arise by asking individuals their religion before participation.

Participants came to the Busara lab for experimental sessions lasting approximately one and a half hours. Each session included up to 25 participants. Sessions were administered by two female Kenyan research assistants, who spoke English and Swahili fluently and were trained in helping participants with comprehension. The experiment was conducted in English.²

Upon arriving at the lab, participants were briefed and then randomly assigned to one of 25 computer workstations with partitions on three sides, so that they were unable to see or speak with the other participants. Within each session, individuals were randomly assigned to one of the three conditions. All of the treatments and measures were implemented on HP TouchSmart 310 desktop computers running Windows 7. Each participant wore headphones and watched the video message on his or her own computer. Participants used the touch screen exclusively to mitigate effects of individual differences in experiences using a mouse and keyboard. All treatments were implemented using z-Tree software (Fischbacher, 2007). At the conclusion of the final questionnaire, participants were debriefed and paid KES 300 in cash.

2.2. Manipulations

The priming literature distinguishes between four types of primes: explicit, implicit, subliminal, and contextual (Shariff et al., 2015). Our study uses explicit messages, as we are most interested in controlling the content of the message. Such control is more difficult with implicit or subliminal primes, which rely on an individual's automatic associations with the concepts presented in the prime. In our case, participants have had little prior exposure to atheism, and thus implicit or subliminal primes would be unlikely to produce strong effects.

The main weakness of explicit priming is its potential to introduce demand effects into the study. By necessity, individuals are more cognizant of the nature of the material they are viewing and are thus more likely to grasp the overall purpose of the experiment. It is therefore possible that individuals would respond by providing the answers they believe the experimenter “wants.” In order to minimize the possibility of this effect, we use an IAT to measure implicit religiosity, which is much less vulnerable to manipulation by the participant (more details below).

The messages used for the study consist of short video presentations (in English) and a writing task in which participants are asked to summarize the main point of the messages. Each participant was randomized into one of three conditions: emotional anti-religious argument, scientific anti-religious argument, and control. Each message is approximately four minutes long and consists of a speaker making arguments against religion, or in the case of the control, speaking about vegetables. The speaker is the same Kenyan narrator in all three videos. Bullet points summarizing the main arguments appear on the screen, accompanying the auditory presentation by the speaker. In the emotional condition, the speaker argues that any God must be very cruel if he allows so many people in the world to suffer, concluding that it is unlikely that God exists. In the scientific condition, the speaker argues that the existence of God is unnecessary to explain the universe given our scientific knowledge. In the control condition, the same speaker argues that it is important to eat vegetables every day. The exact text of the messages is provided in Section A.1 of Appendix. After viewing

the messages, participants are asked to summarize the main arguments in their own words by typing into a text box on their screen and then asked a series of questions to ensure comprehension.

The messages were randomly assigned to conditions within each session, with equal probability for each condition. The setup of the lab included dividers between computers so that participants were unaware of the images shown on the screens of others. Additionally, all participants wore headphones, so they were unaware of what other participants were hearing. All images used in the implicit religiosity measure (described below) were vetted for comprehension by the experimenters and were further validated through debriefing after several pilot sessions.

2.3. Outcome measures

Self-reported religiosity. Our primary outcomes were self-reported and implicit measures of religiosity. Following Shariff et al. (2008), we measure self-reported religiosity with a series of six questions in which participants rate different aspects of their religiosity (belief in God, importance of beliefs and religious traditions) on a scale of one to five. As a primary outcome measure, we analyse a weighted-average index of these responses based on the methodology in Anderson (2008). This index effectively down-weights outcome variables that are highly correlated with each other to avoid redundancy. In the supplemental online material (which may be accessed at <https://doi.org/10.1080/2153599X.2018.1436585>), we also report the results for each of the questions individually and test for joint significance using Seemingly Unrelated Regression (SUR). Thus, in this analysis, we estimate individual regressions for each outcome variable; the critical feature of SUR is that it allows for a test of joint significance across equations. We also present the results for the first latent factor identified using factor analysis on the responses to the six questions.

Implicit religiosity. To control for the possibility of demand effects, we employ a ST-IAT adapted from (but not identical to) the design used by Shariff et al. (2008). We use the ST-IAT to operationalize implicit belief, although we note that strictly speaking it measures associations rather than beliefs. The ST-IAT is a computer-based sorting task that uses response time to measure unconscious associations with a target concept. In each block of the ST-IAT, participants sort three categories of words to the left- and right-hand sides of the screen: synonyms of “real,” synonyms of “imaginary,” and words associated with religion. All language used in the task was vetted for comprehension with the sample population during several pilot sessions. The motivation to use “real” and “imaginary” rather than “true” and “false” for the ST-IAT was to make the three ST-IATs (religion, nature, and cartoon, described below) as similar as possible.

In one block, the participants sort the religious words and synonyms of “real” to the same side. In a second block, the participants sort religious words and synonyms of “imaginary” to the same side. The order in which these two blocks occur is randomized. The ST-IAT design assumes that participants will more quickly sort target words to the side that represents their implicit association with the target.

As a validity check for the ST-IAT, we also administer ST-IATs using alternatively the names of cartoon characters familiar to most Kenyans, and the names of natural landmarks in Kenya. We hypothesize that individuals should be relatively faster to associate cartoons with the concept imaginary, and faster to associate natural landmarks with the concept real.

Using the result of the ST-IAT described above, for each individual we calculate a d-score as

$$d\text{-score} = [\text{mean}(\text{latency}^{\text{real}}) - \text{mean}(\text{latency}^{\text{imaginary}})] / SD^{\text{both}}$$

where $\text{latency}^{\text{real}}$ is the reaction time when associating the concept of interest (religion) with the concept “real,” and $\text{latency}^{\text{imaginary}}$ is the reaction time when associating the concept of interest with the concept “imaginary.” SD^{both} is the pooled standard deviation. Latency (response time) is recorded in

milliseconds; shorter latency indicates a stronger implicit association. Following from the equation above, a lower d-score represents a stronger belief in religious concepts.

Following the recommendations in Greenwald, Nosek, and Banaji (2003), we exclude participants for whom more than 10% of responses are below 300 ms, as well as all responses over 10,000 ms. Individuals who initially respond incorrectly in a trial are required to press the correct response before proceeding to the next trial. We measure latency as the total time from beginning the trial to the entry of a correct response, effectively penalizing incorrect responses with longer latencies (Greenwald et al., 2003). Table S1 in the supplemental online materials reports the mean latencies and accuracy for each of these IATs.

2.3.1. Validity ST-IAT (cartoon characters and nature names)

To determine whether the implicit association test operates as desired in this population, we administered a “Cartoon Character” and “Nature Name” ST-IAT at the very beginning of each experimental session that asks participants to associate cartoon characters and names of animals with “real” and “imaginary.” We reason that on average participants should have “imaginary” associations with cartoon characters, and “real” associations with nature names. If this ST-IAT successfully measures these associations, we can be confident that it is an effective tool for measuring other implicit attitudes.

Other outcomes. We also include a questionnaire on tolerance from the World Values Survey (WVS), which asks for different groups of people, “Of the following groups of people, which would you not like to have as neighbors?” We further include the Cantril ladder to gauge current perceived social status, and expected social status in five years; the negative affect questions from the Positive and Negative Affect Scale (PANAS); and basic demographics questions. Details of these questionnaires are provided in the [Appendix](#).

2.3.2. Session structure

The order of questionnaires, tasks, and messages was as follows.

- (1) Practice IAT.
- (2) Validity ST-IAT (cartoon characters and nature names).
- (3) Video message (four minutes).
- (4) Writing task and comprehension questions.
- (5) Self-reported religiosity questionnaire.
- (6) Practice IAT.
- (7) Religion ST-IAT.
- (8) Cantril ladder.
- (9) WVS tolerance questionnaire.
- (10) PANAS questionnaire.
- (11) Demographics survey.

3. Statistical analysis

3.1. Pre-analysis plan

The study design and all analysis methods were pre-registered before data analysis began (<https://www.socialscienceregistry.org/trials/766>). Our analysis follows this plan strictly, and we point out any additions and deviations below. In addition, we confirm that, for all experiments, we have reported all measures, conditions, and data exclusions. We also confirm that this manuscript includes all studies attempted in this line of work, and that no additional “failed” studies have been “file-drawered.”

3.2. Basic specification

We use ordinary least squares regression to assess the effect of the messages on outcomes. The basic specification is

$$y_i = \beta_0 + \beta_1 T_{R,i} + \beta_2 T_{E,i} + \varepsilon_i, \quad (1)$$

where y_i is the outcome of interest measured at the level of the individual respondent i . $T_{R,i}$ is an indicator taking a value of one if individual i was assigned to the scientific priming condition, and zero otherwise. $T_{E,i}$ is an indicator taking a value of one if individual i was assigned to the emotional priming condition, and zero otherwise. The omitted category is participants assigned to the control condition. ε_i is an idiosyncratic error term. We report heteroskedasticity-robust standard errors. Note that standard errors are not clustered at the session level because randomization occurred within sessions.

Thus, given random assignment to treatment conditions, β_1 identifies the effect of exposure to the scientific messages on the variable interest. β_2 identifies the effect of exposure to the emotional messages on the variable interest. A joint test of β_1 and β_2 identifies the joint effect of the messages.

3.3. Heterogeneous effects

To test for heterogeneous treatment effects, we estimate the following specification:

$$y_i = \beta_0 + \beta_1 T_{R,i} + \beta_2 T_{E,i} + \beta_3 T_{R,i} \times X_i + \beta_4 T_{E,i} \times X_i + \delta X_i + \varepsilon_i. \quad (2)$$

Here, X_i is an indicator for a given dimension of heterogeneity. Given random assignment to treatment conditions, β_3 identifies the heterogeneous effect of exposure to the scientific messages for individuals in group X , and β_4 identifies the heterogeneous effect of exposure to the emotional messages for individuals in group X . We report heteroskedasticity-robust standard errors.

As pre-specified in the pre-analysis plan, the main dimensions of heterogeneity are gender and the z-score of the total number of negative emotions reported by the respondent on the PANAS questionnaire. There is little variation in other demographic measures such as education and age, as we restricted recruitment to current university undergraduates living in Nairobi.

3.4. Controlling for multiple inference

We pre-specified the two main outcomes of interest (self-reported religiosity and implicit religiosity) as separate hypotheses; our main analysis therefore considers them separately and does not adjust for multiple comparisons across them. However, for the sake of completeness, we nevertheless report Family-Wise Error Rate adjusted (FWER) p -values for specification 3.2 across these two outcomes, following Anderson (2008). We also report the p -values for a joint significance test across these outcomes using Seemingly Unrelated Regression (SUR).

4. Results

4.1. Summary statistics and randomization check

Summary statistics for all demographics and outcome variables are detailed in Table 1. The sample includes 183 males and 133 females. All participants were over 18 years of age, with a mean age of 22 and a maximum age of 26. As mentioned above, all participants were enrolled full-time at local Nairobi universities at the time of participation. In some instances, we were unable to administer all questionnaires; thus, for some variables the number of observations is less than 318. Eighty-four percent of the sample self-identified as Christian, suggesting that our recruitment strategy was successful.

Table 2 reports the results of a balance check of key variables across treatment categories: demographics and the results of the two practice IATs that were given before priming. We detect no significant difference in these variables across treatment groups, suggesting that randomization was successful.

Table 1. Summary statistics.

| | Mean | SD | Median | Minimum | Maximum | N |
|---|--------|--------|--------|---------|---------|-----|
| Male | 0.58 | 0.49 | 1.00 | 0.00 | 1.00 | 316 |
| Female | 0.42 | 0.49 | 0.00 | 0.00 | 1.00 | 316 |
| Age | 21.96 | 1.63 | 22.00 | 18.00 | 27.00 | 241 |
| Self-identified Christian | 0.84 | 0.36 | 1.00 | 0.00 | 1.00 | 318 |
| Self-identified Muslim | 0.01 | 0.08 | 0.00 | 0.00 | 1.00 | 294 |
| No religion | 0.02 | 0.13 | 0.00 | 0.00 | 1.00 | 294 |
| Cartoon IAT d-score | −0.14 | 0.52 | −0.20 | −1.39 | 1.29 | 318 |
| Nature IAT d-score | 0.13 | 0.50 | 0.16 | −1.58 | 1.21 | 318 |
| Self-reported religiosity – total score | 24.33 | 4.17 | 25.00 | 6.00 | 30.00 | 318 |
| Implicit religiosity (IAT d-score) | 0.09 | 0.49 | 0.11 | −1.35 | 1.27 | 318 |
| Subjective wellbeing – now (Cantril) | 5.57 | 1.80 | 5.00 | 0.00 | 10.00 | 318 |
| Subjective wellbeing – 5 yr (Cantril) | 8.31 | 1.42 | 9.00 | 1.00 | 10.00 | 318 |
| Tolerance total score (WVS) | −2.11 | 1.22 | −2.00 | −7.00 | 0.00 | 318 |
| PANAS negative affect total | 168.63 | 160.74 | 128.00 | 0.00 | 816.00 | 294 |

Notes: Summary statistics for all demographic and outcome variables. Each row corresponds to a separate characteristic. Male and Female are indicator variables for respondent's gender. Religious categories are indicators for whether a respondent identifies as a member of that religion. Cartoon, nature, and religion IAT results are reported as d-scores for each individual, with higher values corresponding to a faster association with the concept "real." Self-reported religiosity is taken as a total across six questions using a five-point Likert scale, with higher values corresponding to greater religiosity. Both subjective wellbeing scales are 11 point Cantril ladders, with higher values corresponding to greater subjective wellbeing. Total tolerance score is a reverse coding of a total of 10 binary questions described in the supplemental online material. PANAS is a total across 12 questions asking individuals to rate the degree to which they are currently feeling a particular emotion on a 1 to 100 scale. The sample was 318, but in some cases, individuals declined to give a response to specific questions.

4.2. Validity IAT

Before presenting the messages, we test the validity of our ST-IAT setup by running versions with concepts we expect to be more quickly associated with "real" and concepts we expect to be more quickly associated with "imaginary." We use well-known animal and plant names and well-known cartoon characters, respectively. [Figure 1](#) depicts the results of this analysis compared to the full-sample results for the religion ST-IAT, and [Table 3](#) reports the associated statistics. As expected, individuals more quickly associated nature names with the concept "real" than with the

Table 2. Randomization check.

| | (1) Control mean (SD) | (2) Science prime | (3) Emotions prime |
|---------------------------|--------------------------|----------------------|-----------------------|
| Male | 0.58 (0.50) | 0.06 (0.07) | −0.06 (0.07) |
| Female | 0.42 (0.50) | −0.06 (0.07) | 0.06 (0.07) |
| Age | 21.99 (1.68) | −0.09 (0.26) | −0.01 (0.25) |
| Self-identified Christian | 0.81 (0.39) | 0.07 (0.05) | 0.04 (0.05) |
| Self-identified Muslim | 0.01 (0.10) | −0.01 (0.01) | −0.00 (0.01) |
| No religion | 0.03 (0.17) | −0.02 (0.02) | −0.02 (0.02) |
| Cartoon IAT d-score | −0.10 (0.55) | 0.00 (0.07) | −0.12 (0.07) |
| Nature IAT d-score | 0.07 (0.49) | 0.10 (0.07) | 0.08 (0.06) |
| Total observations | 110 | 91 | 117 |

Notes: Results are coefficients and standard errors for OLS regressions of various demographics on indicators for treatment status. Characteristics of interest are listed on the left. Column (1) reports the mean and SD of the control group. Columns (2) and (3) report the coefficients and standard errors from the regression of the characteristic of interest on indicators for the scientific and emotional message, respectively. The last line reports the total number of observations in each condition.

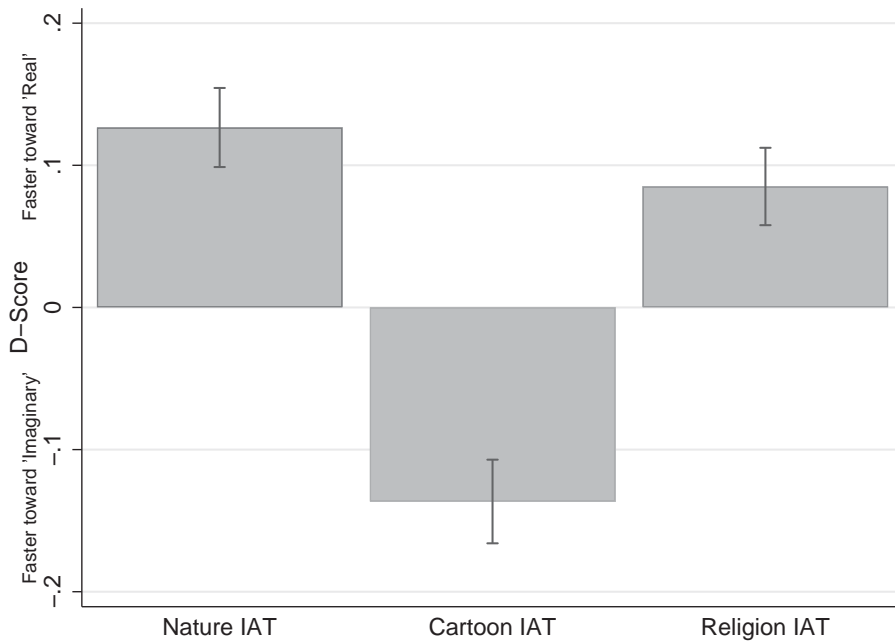


Figure 1. IAT validity check: cartoon versus nature names. Bar graphs show means and one Standard Error of the Mean (SEM) of individual d -scores for the cartoon versus nature single-target IAT. Higher values indicate a faster association between the target concept and the concept “real,” while lower values indicate a faster association between the target concept and the concept “imaginary.”

concept “imaginary,” and associated cartoon characters more quickly with the concept “imaginary” than with the concept “real.” The p -value of the test of equality between the d -scores is <0.01 . This finding is in line with the hypothesis that participants associate items they consider to be real more quickly with the concept “real,” and items they consider to be imaginary more quickly with the concept “imaginary,” suggesting that our ST-IAT is valid.

Figure 1 also shows the results for the ST-IAT for religious words across the entire sample. The corresponding d -score was significantly greater than zero (one-sided t -test, $p < .001$), suggesting that on average participants were more likely to associate religious words with the concept “real” than “imaginary.” This result is consistent with the fact that our sample is largely Christian.

4.3. Treatment effects

Figure 2 displays the effect of treatment of our main self-reported and implicit measures of religiosity in both our full population and when restricted to men only, and Table 3 reports the results of the regressions described in Section 3. More detailed findings are shown in Table S2 in the supplemental online material. Overall, we detect a strong negative effect of exposure to the emotional message on religiosity, especially among men. In the full sample, we detect a decrease of 0.40 SD in self-reported religiosity caused by the emotional message, significant at the 95% level using unadjusted p -values and the 90% level after adjusting for multiple inference. The coefficient on the IAT measure is similarly negative but insignificant. For both measures, there is no significant effect of the scientific message, and the difference between the two messages is not significant. The scientific message does not produce a treatment effect even when restricting the sample to those individuals who passed all comprehension checks, as shown in Table S3 in the supplemental online material. Table S4 in the supplemental online material reports the regression results when restricting the sample to individuals who indicated they were Christians.

Table 3. Treatment effect regressions.

| | All respondents | | | | | Male respondents | | | | |
|--|---------------------|----------------------------|------------------------------|-----------------------------|----------|---------------------|---------------------------|------------------------------|-------------------------------|-----------|
| | (1) Control mean | (2) Science prime | (3) Emotions prime | (4) Difference | (5) N | (6) Control mean | (7) Science prime | (8) Emotions prime | (9) Difference | (10) N |
| Self-reported religiosity index (Anderson) | 0.00 (1.00) | −0.26 (0.17) [0.42] | −0.40** (0.16) [0.07]* | 0.14 (0.19) [0.69] | 318 | −0.07 (1.03) | −0.25 (0.21) [0.73] | −0.43** (0.20) [0.15] | 0.17 (0.22) [0.69] | 185 |
| Implicit religiosity (IAT d-score) | 0.11 (0.49) | 0.00 (0.07) [0.97] | −0.08 (0.06) [0.60] | 0.08 (0.07) [0.56] | 318 | 0.21 (0.45) | −0.09 (0.09) [0.73] | −0.20** (0.08) [0.06]* | 0.11 (0.09) [0.50] | 185 |
| Subjective wellbeing – now (Cantril) | 0.00 (1.00) | 0.35** (0.15) [0.14] | −0.03 (0.13) [0.94] | 0.38** (0.15) [0.10]* | 318 | 0.07 (1.17) | 0.19 (0.22) [0.76] | −0.38** (0.19) [0.16] | 0.57*** (0.19) [0.03]** | 185 |
| Subjective wellbeing – 5 yr (Cantril) | −0.00 (1.00) | 0.25* (0.14) [0.29] | −0.04 (0.13) [0.94] | 0.29** (0.14) [0.16] | 318 | −0.16 (1.12) | 0.31 (0.20) [0.56] | 0.06 (0.17) [0.83] | 0.25 (0.17) [0.50] | 185 |
| Tolerance score (WVS) | 0.00 (1.00) | 0.23 (0.15) [0.42] | 0.14 (0.13) [0.67] | 0.09 (0.15) [0.69] | 318 | 0.18 (0.98) | 0.10 (0.19) [0.84] | 0.09 (0.17) [0.83] | 0.00 (0.19) [0.99] | 185 |
| PANAS negative affect total | 0.00 (1.00) | −0.06 (0.14) [0.88] | 0.21 (0.15) [0.60] | −0.27* (0.16) [0.34] | 294 | 0.00 (1.07) | −0.09 (0.17) [0.84] | 0.23 (0.19) [0.52] | −0.32* (0.17) [0.29] | 175 |
| Joint test (<i>p</i> -value) | | 0.08* | 0.15 | | | | 0.37 | 0.02** | | |

Notes: OLS estimates of the effects of priming on outcome variables. Outcome variables are listed on the left and reported in z-score units with the exception of the IAT results which report individual d-scores. For each outcome variable, we report the coefficients of interest with heteroskedasticity-robust standard errors in parentheses and FWER-adjusted *p*-values in square brackets. Columns (1) through (5) report the results for the full sample. Columns (6) through (10) report the results when restricting to males. Columns (1) and (6) report the mean of the outcome variable among the control group. Columns (2) and (7) report the effect of viewing the scientific message. Columns (3) and (8) report the effect of viewing the emotional message. Columns (4) and (9) report the difference between the effects of the two messages. Columns (5) and (10) report the sample size. *Significance at the 10% level; **significance at the 5% level; ***significance at the 1% level.

Note also that the joint test using SUR is significant for the scientific message, but not the emotional message; this finding results from the insensitivity of SUR to the direction of treatment effects and dilution of the strong effects of the emotional message on self-reported religiosity by small effects on other outcomes.

Our pre-analysis plan pre-specified a heterogeneous analysis by gender. This analysis was largely exploratory in nature, but motivated by the fact that women typically show higher levels of religiosity than men (WVS, 2015), and thus might be more resistant to atheist messages. Indeed, we find that the effects among men are stronger than in the sample as a whole: when restricting the sample to male respondents among male participants, we detect a decrease of 0.30 SD in self-reported religiosity caused by the emotional message, significant at the 95% level using unadjusted *p*-values, and a decrease in implicit religiosity of 0.20 points as measured by the religiosity ST-IAT, significant at the 95% level using unadjusted *p*-values and the 90% level after adjusting for multiple inference. These results suggest men in our sample responded to the anti-religious message, especially when it attacked religion as harmful rather than just untrue. Results for women, and the difference between men and women, are shown in Table S5 in the supplemental online material. Columns (1) and (3) report the treatment effects for women and show that there are no significant effects on religiosity, and the treatment effect of the emotional message on implicit religiosity is significantly larger in men than in women.

Turning to subjective wellbeing, we find that, among men, the emotional message induces a 0.38 SD reduction in self-reported subjective wellbeing as measured by the Cantril ladder, significant at the 95% level. We find no evidence that this pattern extends to individual's expectations of their future wellbeing, as the effects on the five-year Cantril ladder are not significant. Table S5 in the supplemental online material shows that the effect on current wellbeing is significantly larger in men than in women, consistent with the finding that the emotional message reduced religiosity in men but not in women.

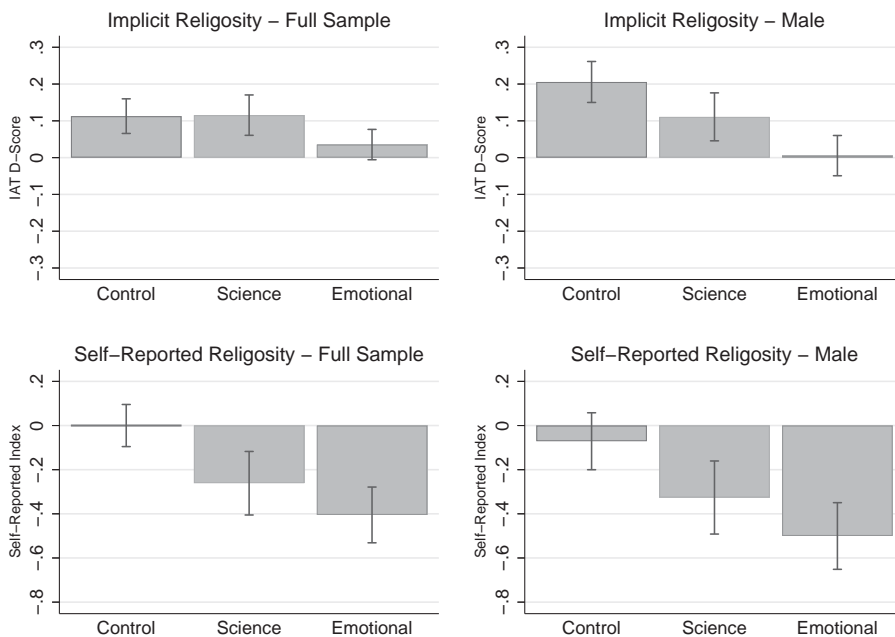


Figure 2. Treatment effects. Top panels: bar graphs show means and one SEM of individual d-scores for the religiosity single-target IAT. Higher values indicate a faster association between the target concept and the concept “real,” while lower values indicate a faster association between the target concept and the concept “imaginary.” Bottom panels: means and one SEM for self-reported religiosity.

Intriguingly, the scientific message actually has the opposite effect on subjective wellbeing in the full population. We detect a 0.35 SD increase in subjective wellbeing as measured by the Cantril ladder, significant at the 99% level, and a 0.25 SD increase in expected future subjective wellbeing, significant at the 90% level. This effect is driven mainly by women, who show a significant 0.60 SD increase in wellbeing after the scientific message. Thus, it appears that the emotional message reduces subjective wellbeing among men, while the scientific message increases it in the whole sample (although only significant in women).

Finally, we detect no significant effect on tolerance, suggesting no change in social preferences due to decreases in religiosity.

5. Discussion

We report on the results of a study in which university students in Kenya were exposed to two different types of atheist messages, following which we measured their religiosity, subjective wellbeing, and tolerance.

5.1. Religiosity

We find that exposure to emotional arguments against religion decreases religiosity. Specifically, we report a strong negative effect of the emotional message on self-reported religiosity of about 0.40 SD. Table S6 in the supplemental material reports the effect for each item included in the self-reported religiosity index. Thus, atheist messages such as those presented by the “New Atheists” may actually affect the intensity of religious belief, at least in the short term. This finding speaks to the broader debate about whether religious beliefs respond to evidence, and suggests that they sometimes do (Levy, 2017; Van Leeuwen, 2014, 2017). It remains to be elucidated what the circumstances are under which this is the case. In addition, it is noteworthy that, in our data, religious belief was changed by arguments in the direction of the arguments, suggesting that “backfire effects,” i.e. a reinforcement of the belief in response to belief-threatening messages, were not of great importance in our setting (Nyhan & Reifler, 2010).

The effect of atheist messages on religiosity is particularly large among males (a pre-specified subgroup analysis), where we find that the emotional message has a strong negative effect on both self-reported (0.43 SD) and implicit (0.20 SD) religiosity. This finding might be understood in light of the fact that women typically show higher levels of religiosity than men (WVS, 2015), and thus might be more resistant to atheist messages. However, we detect no effect of scientific messages on religiosity, in contrast to the results of previous work showing reduced religiosity following exposure to scientific messages (Preston & Epley, 2009; Preston et al., 2013; Shariff et al., 2008).

Our evidence further suggests that religiosity may be more vulnerable to emotional than to scientific appeals, at least in this population. This finding is in line with existing evidence showing that attitudes towards smoking can be affected more with emotional than rational appeals (Smith & De Houwer, 2015). In the study by Smith and De Houwer, this effect was only found among smokers. Drawing the parallel to the present study, this finding suggests that emotional messages may have been more effective than scientific ones because our study population was more religious than those of previous studies. This fact may also explain why scientific explanations do more to reduce religiosity in Western (and more secular) contexts than in ours. In addition, previous evidence suggests that emotional appeals are more effective than rational ones when the target attitudes and behaviors are also emotional (Fabrigar & Petty, 1999); our findings therefore suggest that our participants may conceptualize religion mainly as emotional, rather than scientific.

5.2. Self-reported wellbeing

The effects of the atheism messages on self-reported wellbeing are bi-directional. The emotional atheism message strongly decreases psychological wellbeing among men (0.38 SD). This finding is

in line with previous studies showing that subjective wellbeing is higher among more religious individuals (Deaton & Stone, 2013; Diener et al., 2011), and suggests that atheist arguments may reduce subjective wellbeing. The effect is transitory in that it affects current wellbeing, but not predicted wellbeing five years in the future. However, the effect on current wellbeing is significantly larger in men than in women, consistent with the finding that the emotional message reduced religiosity in men but not in women. One possible mechanism underlying this finding is that the emotional message reminds participants of the suffering in the world, and that this reminder negatively affects wellbeing. The fact that this effect is observed in men but not women may arise from the finding that the message does not affect religiosity in women, but does affect it in men. Thus, women's religiosity, unaffected by the message, may buffer against the negative effect of the emotional message on wellbeing. In contrast, men's religiosity has been reduced by the message and can therefore not protect against the negative wellbeing effects of the message. A caveat to these results is that we observe no overall effect of either message on negative affect, as reported in Tables S7 and S8 in the supplemental online material; they point in the same general direction as the Cantril results, but do not reach significance. One possible reason why the wellbeing results are not stronger may be that the self-report measures of religion that followed the messages mitigated the negative effect of the messages on wellbeing (Jonas & Fischer, 2006).

The scientific atheism message strongly increases wellbeing among both men and women by a similar magnitude (0.35 SD). Thus, different arguments for atheism may differentially affect psychological wellbeing, and these effects may be heterogeneous by gender; future work should test the robustness of this finding. A possible mechanism underlying the positive wellbeing affect of the scientific message may be that it makes respondents feel liberated, while still providing an explanation for events in the world; however, this explanation is at odds with the fact that the scientific message does not in fact reduce religiosity. An alternative account is that the increase in wellbeing induced by the scientific message may increase participants' perceived orderliness and control of events in the world, and thereby raise wellbeing even in the absence of changes in religiosity (Rutjens, van Harreveld, & van der Pligt, 2010; Rutjens, van Harreveld, van der Pligt, Kreemers, & Noordewier, 2013). Future work might test these hypotheses.

5.3. Tolerance

Finally, we find no evidence that a decrease in religiosity is associated with either increases or decreases in tolerance, as reported in Table S9 of the supplemental online material. This finding is at odds with previous studies that find religious messages cause racial bias and strong in-group preferences (Johnson et al., 2010; LaBouff et al., 2012; Ramsay et al., 2014). We note, however, that it may represent an order effect as the tolerance measures were presented after those for religiosity and self-reported wellbeing.

5.4. Future work

Future work might assess whether atheist messages similar to those used in the present study impact other outcomes beyond religiosity, wellbeing, and tolerance. A number of previous studies have posed similar questions for religious messages, showing that they increase honesty (Randolph-Seng, 2007), punishment of unfairness (McKay, Efferson, Whitehouse, & Fehr, 2010), self-control (Laurin, Kay, & Fitzsimons, 2012; Rounding, Lee, Jacobson, & Ji, 2012), risk-taking (Chan, Tong, & Tan, 2014; Kuper, Laurin, & Levav, 2015), and political participation (McClendon & Riedl, 2015), although we note that some of these effects have significance levels that would be classified as "suggestive" in the new terminology of Benjamin et al. (2018). In light of this evidence connecting religious messages to increases in pro-social behavior, we deem it important to evaluate whether anti-religious messages have the opposite effect. Additionally, one might look for effects of atheist messages on in-group and out-group preferences, honesty, and risk-taking.

In doing so, future studies might take into account the possibility that what was being primed in these earlier studies on religious messages and prosocial behavior, and possibly in the present one, is not religion or atheism narrowly, but the presence of a watchful authority (Norenzayan & Shariff, 2008; Randolph-Seng, 2007; Yilmaz & Bahçekapili, 2016); it will be interesting to ask if religious messages have effects over and above those referring to a watchful authority.

A further topic for exploration concerns the utility of the ST-IAT to measure implicit religiosity. We use here an operational definition of implicit religiosity based on implicit associations as measured by the ST-IAT. However, the effect of our messages on the ST-IAT could potentially reflect variables other than personally held beliefs, such as positive versus negative valence or culturally determined associations. Future work might attempt to distinguish between these different possibilities.

Together, the answers to these questions may contribute to understanding whether the message promulgated by thinkers like the New Atheists is socially beneficial or harmful. By implication, these results would give us better insight into the role of religion in society, both in terms of its benefits and harms.

Notes

1. USD values are calculated at purchasing power parity, using the 2014 World Bank PPP estimate for private consumption in Kenya: 0.024.
2. Restricting participation to university students ensured English comprehension. Although Kenyan universities do not typically require an official test of English proficiency, matriculating students are expected to be proficient in written and spoken English, and much of the instruction is in English. Additionally, Busara has confirmed through previous studies that the vast majority of Kenyan university students are highly proficient in English.

Acknowledgments

We are grateful to the study participants for generously giving their time, and to Ara Norenzayan, Azim Shariff, and conference participants at The Symposium on Economic Experiments in Developing Countries for comments and discussion.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Johannes Haushofer  <http://orcid.org/0000-0002-9984-9113>

References

- Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training projects. *Journal of the American Statistical Association*, 103 (484), 1481–1495. doi:10.1198/016214508000000841
- Ano, G. G., & Vasconcelles, E. B. (2005). Religious coping and psychological adjustment to stress: A meta-analysis. *Journal of Clinical Psychology*, 61(4), 461–480.
- Bassett, J. F., & Connelly, J. N. (2011). Terror management and reactions to undocumented immigrants: Mortality salience increases aversion to culturally dissimilar others. *The Journal of Social Psychology*, 151(2), 117–120.
- Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E. -J., Berk, R., ... Johnson, V. E. (2018). Redefine statistical significance. *Nature Human Behaviour*, 2(1), 6–10. doi:10.1038/s41562-017-0189-z
- Bloom, P. (2007). Religion is natural. *Developmental Science*, 10(1), 147–151. doi:10.1111/j.1467-7687.2007.00577.x
- Chan, K. Q., Tong, E. M. W., & Tan, Y. L. (2014). Taking a leap of faith: Reminders of God lead to greater risk taking. *Social Psychological and Personality Science*, 5(8), 901–909. doi:10.1177/1948550614537309
- Deaton, A., & Stone, A. A. (2013). Two happiness puzzles. *American Economic Review*, 103(3), 591–597. doi:10.1257/aer.103.3.591

- Diener, E., Tay, L., & Myers, D. G. (2011). The religion paradox: If religion makes people happy, why are so many dropping out? *Journal of Personality and Social Psychology*, 101(6), 1278–1290. doi:10.1037/a0024402
- Egloff, B., & Schmukle, S. C. (2002). Predictive validity of an Implicit Association Test for assessing anxiety. *Journal of Personality and Social Psychology*, 83(6), 1441–1455.
- Fabrigar, L. R., & Petty, R. E. (1999). The role of the affective and cognitive bases of attitudes in susceptibility to affectively and cognitively based persuasion. *Personality and Social Psychology Bulletin*, 25(3), 363–381. doi:10.1177/0146167299025003008
- Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2), 171–178. doi:10.1007/s10683-006-9159-4
- Greenberg, J., Pyszczynski, T., Solomon, S., Rosenblatt, A., Veeder, M., Kirkland, S., & Lyon, D. (1990). Evidence for terror management theory II: The effects of mortality salience on reactions to those who threaten or bolster the cultural worldview. *Journal of Personality and Social Psychology*, 58(2), 308–318. doi:10.1037/0022-3514.58.2.308
- Greenberg, J., Simon, L., Pyszczynski, T., Solomon, S., & Chatel, D. (1992). Terror management and tolerance: Does mortality salience always intensify negative reactions to others who threaten one's worldview? *Journal of Personality and Social Psychology*, 63(2), 212–220.
- Greenberg, J., Solomon, S., & Pyszczynski, T. (1997). Terror management theory of self-esteem and cultural worldviews: Empirical assessments and conceptual refinements. *Advances in Experimental Social Psychology*, 29, 61–139. doi:10.1016/S0065-2601(08)60016-7
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, 85(2), 197–216. doi:10.1037/0022-3514.85.2.197
- Johnson, M., Rowatt, W., & LaBouff, J. P. (2010). Priming Christian religious concepts increases racial prejudice. *Social Psychological and Personality Science*, 1(2), 119–126.
- Jonas, E., & Fischer, P. (2006). Terror management and religion: Evidence that intrinsic religiousness mitigates worldview defense following mortality salience. *Journal of Personality and Social Psychology*, 91(3), 553–567. doi:10.1037/0022-3514.91.3.553
- Kupor, D. M., Laurin, K., & Levav, J. (2015). Anticipating divine protection? Reminders of God can increase nonmoral risk taking. *Psychological Science*, 26(4), 374–384. doi:10.1177/0956797614563108
- LaBouff, J. P., Rowatt, W. C., Johnson, M. K., & Finkle, C. (2012). Differences in attitudes toward outgroups in religious and nonreligious contexts in a multinational sample: A situational context priming study. *The International Journal for the Psychology of Religion*, 22(1), 1–9. doi:10.1080/10508619.2012.634778
- 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(1), 4–21. doi:10.1037/a0025971
- Levy, N. (2017). Religious beliefs are factual beliefs: Content does not correlate with context sensitivity. *Cognition*, 161, 109–116. doi:10.1016/j.cognition.2017.01.012
- Major, B., Kaiser, C. R., O'Brien, L. T., & McCoy, S. K. (2007). Perceived discrimination as worldview threat or worldview confirmation: Implications for self-esteem. *Journal of Personality and Social Psychology*, 92(6), 1068–1086.
- McClendon, G., & Riedl, R. B. (2015). Religion as a stimulant of political participation: Experimental evidence from Nairobi, Kenya. *The Journal of Politics*, 77(4), 1045–1057. doi:10.1086/682717
- McKay, R., Efferson, C., Whitehouse, H., & Fehr, E. (2010). Wrath of God: Religious primes and punishment. *Proceedings of the Royal Society of London B: Biological Sciences*. doi:10.1098/rspb.2010.2125
- Norenzayan, A., & Shariff, A. F. (2008). The origin and evolution of religious prosociality. *Science*, 322(5898), 58–62. doi:10.1126/science.1158757
- Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2005). Understanding and using the Implicit Association Test: II. Method variables and construct validity. *Personality and Social Psychology Bulletin*, 31(2), 166–180. doi:10.1177/0146167204271418
- Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, 32(2), 303–330. doi:10.1007/s11109-010-9112-2
- Preston, J., & Epley, N. (2009). Science and God: An automatic opposition between ultimate explanations. *Journal of Experimental Social Psychology*, 45(1), 238–241. doi:10.1016/j.jesp.2008.07.013
- Preston, J. L., Ritter, R. S., & Hepler, J. (2013). Neuroscience and the soul: Competing explanations for the human experience. *Cognition*, 127(1), 31–37. doi:10.1016/j.cognition.2012.12.003
- Ramsay, J. E., Pang, J. S., Shen, M. J., & Rowatt, W. C. (2014). Rethinking value violation: Priming religion increases prejudice in Singaporean Christians and Buddhists. *International Journal for the Psychology of Religion*, 24(1), 1–15. doi:10.1080/10508619.2012.761525
- Randolph-Seng, B. (2007). Honesty: One effect of primed religious representations. *International Journal for the Psychology of Religion*, 1(1), 303–315.
- Rounding, K., Lee, A., Jacobson, J. A., & Ji, L. J. (2012). Religion replenishes self-control. *Psychological Science*, 23(6), 635–642. doi:10.1177/0956797611431987
- Rutjens, B. T., van Harreveld, F., van der Pligt, J. (2010). Yes we can: Belief in progress as compensatory control. *Social Psychological and Personality Science*, 1(3), 246–252. doi:10.1177/1948550610361782

- Rutjens, B. T., van Harreveld, F., van der Pligt, J., Kreemers, L. M., & Noordewier, M. K. (2013). Steps, stages, and structure: Finding compensatory order in scientific theories. *Journal of Experimental Psychology: General*, 142(2), 313–318. doi:10.1037/a0028716
- Schimel, J., Hayes, J., Williams, T., & Jahrig, J. (2007). Is death really the worm at the core? Converging evidence that worldview threat increases death-thought accessibility. *Journal of Personality and Social Psychology*, 92(5), 789–803.
- Shariff, A., Cohen, A., & Norenzayan, A. (2008). The Devil's advocate: Secular arguments diminish both implicit and explicit religious belief. *Journal of Cognition and Culture*, 8, 417–423. doi:10.1163/156853708X358245
- Shariff, A. F., Willard, A. K., Andersen, T., & Norenzayan, A. (2015). Religious priming: A meta-analysis with a focus on prosociality. *Personality and Social Psychology Review*. doi:10.1177/1088868314568811
- Smith, C. T., & De Houwer, J. (2015). Hooked on a feeling: Affective anti-smoking messages are more effective than cognitive messages at changing implicit evaluations of smoking. *Frontiers in Psychology*, 6, 1488. doi:10.3389/fpsyg.2015.01488
- Van Leeuwen, N. (2014). Religious credence is not factual belief. *Cognition*, 133(3), 698–715. doi:10.1016/j.cognition.2014.08.015
- Van Leeuwen, N. (2017). Do religious “beliefs” respond to evidence? *Philosophical Explorations*, 20(S1), 52–72. doi:10.1080/13869795.2017.1287294
- WVS. (2015). World Values Survey (WVS) wave 6 (2010–2014) (Aggregated Documentation). v.20150418. World Values Survey Association Aggregate. File Producer: Asep/JDS, Madrid, Spain. Retrieved from <http://www.worldvaluessurvey.org/WVSDocumentationWV6.jsp>
- 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. doi:10.1016/j.evolhumbehav.2015.09.005

Appendix. Experiment details

A.1. Messages

A.1.1. Scientific atheism message

- Most people think that you need God to explain the existence of life, but they are wrong.
- Science offers us an explanation of how the complexity of life arose out of simplicity without God.
- The theory that God exists does not offer a good explanation for anything.
- The origin of life was a chemical event, whereby a simple molecule able to replicate itself and pass down its structure (such as DNA) came into existence.
- This strikes many people as improbable, but scientists invoke large numbers to explain how life arose despite the unlikelihood.
- With so many planets in the universe, life still would have emerged on one billion planets.
- Once life came into existence, Charles Darwin's theory of evolution by natural selection explains how complex life forms arose step by step without a God.
- Millions of mutations happen in life forms all the time, but only those that make an organism better able to survive are passed down.
- Over millions of years, these beneficial mutations add up, and organisms become more and more complex and well suited to their environment.
- The alternative that life was designed by God is unnecessary, and it is also highly improbable.
- Any designer like God capable of designing something really complex like life, has to be really complex himself.
- The existence of God would raise a bigger mystery than it solves because it does not tell us where God came from, so God is not a good explanation for the universe.

A.1.2. Emotional atheism message

- The widespread pain and suffering in the world is very strong evidence that God does not exist.
- 9,000,000 children die every year before they reach the age of five. That is 24,000 children a day, 1000 per hour, 17 per minute.
- Any God that would allow children by the millions to suffer and die in this way and for their parents to grieve in this way is either powerless or evil.
- According to many Christians, millions of people will go to Hell simply because they were born into the wrong religion.
- If God created these people, then he is responsible for the fact that they were brought up with a different religion and are ignorant of Christianity.

- God created a world in which they would not be Christians, and then he created a penalty for this ignorance that he created, so he must be evil.
- In the Christian viewpoint, any murderer can spend his life raping and torturing children, but will be forgiven by God if he converts to Christianity.
- Christians use a double standard to exonerate their vision of God.
- Given everything that God does not do to help people, believing in God is morally reprehensible.
- If God exists and is good, he would not have given us a book (the Bible) that defends slavery and tells us to kill people for imaginary crimes like witchcraft.
- However it is not just the Christian God that is immoral and evil – any religion that justifies this killing is immoral and evil.
- It is clear that either God does not exist or he is evil.

A.1.3. Control message

- You need to eat vegetables everyday because you simply cannot find another food group that is as perfectly matched to our everyday human needs.
- The nature of vegetables and the nature of human health are matched up in a way that simply cannot be duplicated by other food groups.
- To begin with, vegetables as a group are so low in calories that it is very difficult to gain weight even if you overeat them.
- On average, there are only 50 calories or less per cup from most of the healthiest vegetables.
- That amount is astonishingly low, even when you compare it to other food groups.
- The uniquely low-calorie nature of vegetables as a group means that you can eat a lot of them and not have to worry about the calories.
- Optimal nutrition is another reason that vegetables are important on a daily basis because you need a supply of vitamins every day.
- When considered as a group, vegetables are unusually rich sources for a full mixture of water-soluble vitamins.
- When it comes to vegetables, there is also their abundance of phytonutrients to consider.
- Finally is the pleasure of chewing and amazing digestive benefits that come from the high-fiber content of vegetables.
- Food cannot move through our digestive tract in a healthy way unless it is fiber-rich, and vegetables are some of the very richest sources of fiber that exist.
- It is clear that you should eat a lot of vegetables every day.

A.2. Additional outcome measures

A.2.1. Cantril ladder

To assess how religion may influence an individual's sense of status, we present participants with a Cantril ladder, which asks participants where they would rate their current life compared to the best possible life they can imagine for themselves, both now and in five years, by indicating where they stand (will stand) on a picture of a ladder with steps labeled 0 to 10. The question asked is: "Assume that this ladder is a way of picturing your life. The top of the ladder represents the best possible life for you. The bottom rung of the ladder represents the worst possible life for you. Indicate where on the ladder you stand {right now; five years from now}."

We calculate z-scores for responses to both the present date and five-year Cantril ladder question.

Tolerance. We adapt a series of questions from the World Values Survey asking participants "Of the following groups of people, which would you not like to have as neighbors?"

- Drug addicts
- People of a different race
- People who have AIDS
- Immigrants/foreign workers
- Homosexuals
- People of a different religion
- Atheists/people with no religion
- Heavy drinkers
- Unmarried couples living together
- People who speak a different language

We analyse individual responses and test joint significant with SUR regression; in addition, we calculate both aggregate and weighted average indexes across all responses.

A.2.2. Positive and Negative Affect Schedule (PANAS)

To assess the possible effect of the messages on affect, individuals complete questions from the PANAS evaluating negative affect after viewing the messages. On a scale of 1 (not at all) to 100 (very much), participants are asked “How much do you feel ...”

- “Distressed, at this moment?”
- “Upset, at this moment?”
- “Guilty, at this moment?”
- “Ashamed, at this moment?”
- “Hostile, at this moment?”
- “Irritable, at this moment?”
- “Nervous, at this moment?”
- “Jittery, at this moment?”
- “Scared, at this moment?”
- “Afraid, at this moment?”
- “Frustrated, at this moment?”
- “Stressed, at this moment?”

We evaluate whether priming impacted affect by analysing individual responses and testing for joint significance with SUR regression; in addition, we analyse the summary indices of this measure.

A.2.3. Demographics

We ask individuals to report their gender, age, and religious affiliation.