Belief in Belief

Even Atheists in Secular Countries Show Intuitive Preferences Favoring Religious Belief

Will M. Gervais [*] Ryan T. McKay [†] Jazmin L. Brown-Iannuzzi [‡]

Robert M. Ross [§] Gordon Pennycook [¶] Jonathan Jong [‖] Jonathan A. Lanman [**]

February 2025
MANUSCRIPT IN PRESS, *PNAS*

Contents

[*]Centre for Culture and Evolution, Brunel University London, will.gervais@brunel.ac.uk
[†]Royal Holloway University of London
[‡]University of Virginia
[§]Macquarie University
[¶]Cornell University
[‖]Coventry University
[**]Queen's University Belfast

Abstract

We find evidence of *belief in belief* – intuitive preferences for religious belief over atheism, even among atheist participants – across 8 comparatively secular countries. Religion is a cross-cultural human universal, yet explicit markers of religiosity have rapidly waned in large parts of the world in recent decades. We explored whether intuitive religious influence lingers, even among nonbelievers in largely secular societies. We adapted a classic experimental philosophy task to test for this intuitive belief in belief among people in eight comparatively nonreligious countries: Canada, China, Czechia, Japan, the Netherlands, Sweden, the UK, and Viet Nam (total $N = 3804$). Our analyses revealed strong evidence that (1) people intuitively favor religious belief over atheism and that (2) this pattern was not moderated by participants' own self-reported atheism. Indeed, (3) even atheists in relatively secular societies intuitively prefer belief to atheism. These inferences were robust across different analytic strategies, and across other measures of individual differences in religiosity and religious instruction. Although explicit religious belief has rapidly declined in these countries, it's possible that belief in belief may still persist. These results speak to the complex psychological and cultural dynamics of secularization.

Keywords: atheism; religion; intuitions; culture; evolution; dual inheritance theory

Significance Statement

Religion is a cross-cultural human universal, and religions may have been instrumental in the cultural evolution of widespread cooperation and prosociality. Nonetheless, religiosity has rapidly declined in some parts of the world over just a handful of decades. We tested whether longstanding religious influence intuitively lingers, even in overtly secular and nonreligious societies. Using a classic experimental philosophy task, we found that even atheists in nonreligious societies show evidence of intuitive preferences for religious belief over atheism. This is the first compelling cross-cultural experimental evidence for intuitive preferences for religion among nonbelievers – a hypothesized phenomenon that philosopher Daniel Dennett dubbed *belief in belief*.

Introduction

> "Many people believe in *belief in God*. . . .[they] are sure that *belief in God* exists (and who could doubt that?) and they think this is a good state of affairs, something to be strongly encouraged and fostered wherever possible: If only *belief in God* were more widespread!"
>
> ~ Daniel C. Dennett, *Breaking the Spell*

Might even atheists in highly secular countries show intuitive preferences favoring religious belief? The possibility that nonbelievers might harbor pro-religious preferences – dubbed *belief in belief* by Dennett (1942-2024) – has not received much direct scholarly attention since it was initially proposed nearly 20 years ago (1). However, the putative existence of intuitive belief in belief may have important implications for the cultural and evolutionary study of religion.

In this paper, we test for the existence of intuitive belief in belief, and we leverage cultural evolutionary theory, experimental philosophy, and cross-cultural psychological methods to do so. Cultural evolution gives us a compelling theoretical framework for understanding *why* belief in belief might exist and persist; experimental philosophy tells us *how* we might look for belief in belief and furnishes us with rigorous tools for assessing intuitive preferences; cultural psychology lets us focus on questions of *where* we might draw samples for the most compelling tests of belief in belief worldwide.

Cultural Evolution: *Why* Belief in Belief

Within a few short generations, societies that were typified by devout belief for centuries-to-millennia are now characterized instead by religious disaffiliation and disbelief (2, 3). But because religions may have been key contributors to the expansion of human prosociality (4–6), and have thus played a key role in shaping our cooperative and moral intuitions and norms (7–9), pro-religious intuitions might persist even as overt religiosity wanes (10). Tensions between these two sets of cultural evolutionary forces – those linking religion and morality and those driving secularization in some parts of the world, respectively – may provide a theoretical framework for understanding why some people might hold the intuitive belief that, whether or not one personally believes in a god, belief in god is nonetheless good – belief in belief (1).

Religious Prosociality and Moral Intuitions

Humans display a much higher degree of prosocial and cooperative behavior than do our closest primate relatives. Our surprising prosociality is a fairly recent cultural evolutionary innovation, and varies considerably across cultures (11–13). How did humans become so cooperative over the past 12 millennia or so, and why do

different human groups today show such disparate levels of cooperation?

Religions may help explain both puzzles, at least according to a prominent view of religious prosociality (4–7). In this view, certain religious beliefs and norms have helped cement cooperation within groups that held them[i]. Specifically, beliefs in big moralizing gods can help curb selfish impulses and mitigate freerider problems that ordinarily undermine cooperation in big groups. Successful religions also incorporate synchronous ritual practices (15), credible behavioural displays (16, 17), and explicit moral teachings (9).

Religious beliefs, norms, and practices can all interact to promote cooperation and group cohesion, but this has not been uniform across religious traditions. Religions that lacked the beliefs, practices, and norms that promote cooperation and cohesion have been plentiful throughout history – but haven't tended to last or spread (5). Our current global religious landscape is largely one defined by the religions that "won" the cultural evolutionary arms race to attract adherents, build successful coalitions, and outcompete rival religions (10). Most successful religions include explicit moral teachings that are sanctioned by a deity (or deities) who is concerned about them, and able to police adherence to them. As a handful of moralizing religions spread throughout the globe, the lines between religion and morality have become blurred in the minds of adherents (9). This seems to have led many people to – explicitly or more intuitively – perceive religion as essential to, *if not synonymous with*, morality. People regularly report, for example, that morality would be impossible without belief in a god, or that religion is a necessary ingredient in a child's moral upbringing (18). One consequence of this is a deep-seated intuition conflating a lack of religious belief with a lack of morality (19–22).

What, then, to make of regional declines in religiosity?

Stages of Secularization

Religions may have been instrumental in the upscaling of human cooperation, but they are also on the decline in large parts of the world (2, 3). If religions helped build our current interconnected and cooperative world, why are they fading away in some pockets of the globe?

One possibility is that religions' cooperative successes may have indirectly led to their (partial and localized) downfall. Religions built large, cooperative societies that in turn created large institutions to cement even greater degrees of cooperation. In the process, these societies tended to exploit human and natural capital to extract wealth to power their expansion and stability – with eventually detrimental effects for religion itself. As such societies come to experience greater degrees of existential security – health, wealth, prosperity, education, and the like – it seems that people are less motivated by religion (2, 3, 23, 24). This isn't to say that people rapidly abandon their beliefs. Instead people at first tend to be just a bit less fervent

[i] see, though, (14)

and public in their religion. Public displays of religiosity decline, and those remaining outward displays of religious commitment may mask private disbelief (24, 25).

Public religious apathy may in turn breed religious disbelief among subsequent generations of cultural learners. Learners attend to the credible behavioral cues that others give to mark the sincerity of their beliefs (16). If learners observe credibility enhancing displays (CREDs) of others' faith in a given god, they are more likely to believe as well (26–29). Without consistent cues to believe in any given god, atheism may be a natural result (10). Consistent with this, a lack of credibility enhancing religious displays emerged as a consistent predictor of atheism in nationally representative samples in the USA, Czechia, and Slovakia (30, 31).

This suggests a two-stage process. First, prosperity – if relatively equitably channeled towards stability and existential security – breeds religious complacency and public disengagement among one generation. Second, the next generation of cultural learners, deprived of consistent credibility-enhancing displays supporting specific faiths, abandons religious belief more completely. This two-stage process can occur rapidly, as evidenced by declines of religiosity in much of Western Europe over the past several decade (2, 3, 32).

But how psychologically pervasive are these effects – is rapid explicit secularization mirrored by a concomitant degrading of intuitions favoring religion?

Secularism, Fast and Slow

Attendance at religious services can drop precipitously over a fairly short period of time. For example, a 2018-2019 Pew survey showed that 17% of Americans never attend religious services, a more than 50% increase in abstention from just a decade prior. Underlying beliefs might take more time to change. In general, declines in belief appear to be greater across successive cohorts than within the same cohorts over time (24, 33). Subsequent generations believe less than those that precede them, rather than individuals within each generation becoming less religious over time. This fits the previously-sketched account well: security drives down attendance, which in turn drives down belief as subsequent cultural generations witness fewer religious CREDs. Within just a few decades, populations can shift from mostly fervent and public belief, to societies characterized by little public religious attendance and diminished private belief. However, this does not mean that religion no longer has any influence.

In cultural evolution, initial causes can have rippling consequences over time (34, 35). For example, Schulz, Henrich, and colleagues have argued that the medieval Church initiated a series of shifts in kinship and cooperative norms that led to a cascade of effects that can explain large-scale patterns in cross-cultural psychology today (36). Might religious influence still be evident, even in overtly nonreligious societies? Religious supremacy in the moral and cooperative spheres over centuries-to-millennia could, we hypothesize,

leave psychological and cultural traces that can't be erased in just a few decades. With this in mind, it's plausible that some latent religious influence might still be felt, even in highly secular societies. If this is the case, then we might expect that people in highly secular societies nonetheless show some favoritism for religion, albeit at an intuitive level. In other words, rapid overt secularism may overlie some latent pro-religious intuitions. Explicit secularism can be rapid, but subtle intuitive preferences for religion may linger even as attendance at religious services and belief in gods fade.

Consistent with this possibility, many people in religious countries show an extreme degree of moral distrust for atheists – people intuitively assume that perpetrators of immoral deeds like incest, cannibalism, necrobestiality, and serial murder must be atheists (21). This pattern persists even among participants who themselves are atheists (21), and a 13-country investigation found that although the magnitude of moral distrust of atheists is predictably stronger in more religious countries, it remains evident in highly secular countries as well – again, even among atheist participants (20). Now, associating atheists with severe immorality up to serial murder and cannibalism aligns with a belief in belief account (e.g., associating someone's identity with incest or necrobestiality is inconsistent with intuitively favoring their beliefs), but is not synonymous with belief in belief across the board. Associating atheists with specific and extreme moral violations is not the same as more generally finding belief more laudable than atheism, and the use of specific immorality vignettes in prior work limits their applicability to the present work. Extreme moral distrust of atheists may be just one manifestation of a broader pattern whereby even atheists in secular societies nonetheless show some intuitive preference for religion, but methodological reliance on vignettes depicting specific moral misdeeds does not itself demonstrate intuitive general preferences for religious belief. Belief in belief is about more than just suspecting that serial killers and cannibals don't believe in God; it reflects a *general* sense that the world would be better off with more faith and less atheism. Such a general intuitive preference can be ably measured using classic experimental philosophy methods.

Experimental Philosophy: *How* to Measure Intuitive Belief in Belief

Our focal measure of intuitive preference for religion, *vis a vis* atheism, used a classic experimental philosophy task known as the 'side effect effect' or 'Knobe effect' (37). In the original version, participants are given a vignette about a CEO mulling a new policy for their company. An advisor recommends a new policy that will make the company money, and notes that it will also either help or hurt the environment (varying between participants as an experimental manipulation). The CEO adopts the policy, money is made, and the environment is either helped or harmed. Participants are asked to rate whether the CEO intentionally harmed/helped the environment. When the outcome is negative (harming the environment), participants

reliably rate it as more intentionally caused than when the outcome is positive (helping the environment). Over the years, work on the Knobe effect has focused both on mechanisms and related outcomes. Is it that *specifically moral* bads are seen as more intentionally caused, or is it that mere negative valence yields inferences of intent? What is the involvement of theory of mind (38)? Aside from inferences about intentionality, can the moral valence of actions lead to inferences about the knowledge of agents involved (39)?

These precise details of how and why the side-effect effect works are beyond the scope of our work, but the general task can be useful for studying potentially subtle intuitions about the (perhaps moral) goodness or badness of religious belief and atheism. Classic interpretations of the Knobe effect suggest that morally negative outcomes are seen as more intentionally caused, meaning that inferences can be drawn about intuitive preferences for competing options by seeing which sorts of side effects participants rate as being intentionally caused. Outcomes rated as intentionally caused, by inference, are intuitively viewed as more negative than outcomes rated as more incidental. In short: the degree to which an action's side effect is rated as intentionally caused can be used to indirectly gauge intuitions participants have about the moral goodness or badness of different outcomes.

We capitalized on this effect by designing a vignette in which the side effect would be that more people either 1) become atheists, or 2) come to believe in God. We then asked participants to rate whether or not the side effect (increasing the number of atheists or increasing the number of believers) was intentionally caused. By comparing intentionality judgments across conditions, we could form inferences about people's intuitive preferences for religious belief, relative to atheism.

Cultural Psychology: *Where* to Look for Belief in Belief

Belief in belief implies that there exist nonbelievers who nonetheless hold intuitive preferences for religious belief (vis a vis atheism) – that some atheists find something intrinsically laudable about faith that they themselves don't have. Presumably, intuitive preferences for religion among atheists might be easiest to find in nominally religious countries, with overt and explicit norms promoting religion. A far more interesting and theoretically compelling possibility is that, consistent with the cultural evolutionary theorizing above, *even atheists in currently highly secular locales* might nonetheless exhibit belief in belief. In this light, the strongest test of belief in belief would selectively recruit people from locales with little overt pro-religious influence. To that end, we embedded our belief in belief measure within a survey being fielded across some of the most secular countries on earth: Canada, China, Czechia, Japan, the Netherlands, Sweden, the UK, and Viet Nam. In these countries, religious belief and practices are comparatively low, and all show low levels of

explicit pro-religious (or anti-atheistic) responding in large-scale polls (3, 18). This sampling frame offers a conservative and general test of our core hypothesis, highlighting secular societies with different cultural, economic, and political backgrounds.

Overview

We investigated the possible existence of belief in belief (e.g., pro-religious intuitive preferences) in secular societies, with a special interest in whether even atheists might hold them. To this end, we deployed an experimental philosophy task measuring intuitive preferences across 8 increasingly secular and nonreligious countries. We tested two primary hypotheses across a series of nested analyses:

1. That people would tend to intuitively prefer religious belief over atheism.
2. That an intuitive preference for religion over atheism would even be present among atheist participants.

To establish robustness of findings, we considered a range of different operationalizations and analytic strategies.

Results

Modeling and Inferential Approach

Our data were nested within our eight countries, and we used a series of hierarchical Bayesian models (40–42) to account for country-level dependencies. Three nested models allowed us to quantify evidence pertaining to our primary hypotheses. Model results are summarized with point estimates depicting the most credible parameter estimate (posterior mode), and we indexed estimate precision and uncertainty with the 89% highest posterior density (HPDI), the region in which the most credible 89% of estimates lie. We note that an 89% interval is no more arbitrary than any other threshold value, and we've resisted the urge to use a familiar-seeming 95% interval width to remind our readers that our Bayesian credible intervals are neither statistical significance tests nor frequentist confidence intervals (although people seem to mis-intuit that frequentist intervals have Bayesian properties (43, 44)). Our 89% credible intervals align with best practices in Bayesian inference, which allow for flexibility in selecting intervals that best represent the posterior distribution. As emphasized in recent literature (41, 45), Bayesian methods enable the use of interval widths tailored to the data and research questions, rather than adhering to traditional frequentist conventions like 95%. Our decision to use an 89% credible interval – following McElreath's prime directive (41) – provides a flexible balance between a precise representation of uncertainty and the broader range of posterior probabilities.

More broadly, we remind readers that the credible intervals primarily illustrate the precision and uncertainty in our estimates, and we urge readers not to use these credible intervals to make dichotomous statistical-significance-like judgments about which intervals overlap which values in a plot. Instead, we urge readers to focus their binary inferential instincts on the posterior probabilities that directly index the probability of observed differences, contingent on data and modelling assumptions. In sum: our most credible parameter estimates are reflected in point estimates, estimate precision and uncertainty are reflected in the posterior density intervals, and the likelihood of observed differences are indexed by the posterior probabilities – these values reflect the probability of specific directionally hypothesized pairwise differences, given data and modeling assumptions.

We asked of our data three key questions: Do people intuitively prefer belief to atheism? Do intuitive preferences for belief interact with participant atheism? Do even atheists in secular countries show evidence of intuitive belief in belief? To do so, we took a model comparative approach across three nested hierarchical (i.e., mixed effects) models, accounting for country-level clustering, and calculated Bayes factors for the degree to which evidence favors one model or the other. Models used the brms package in R (46), including its in-built Bayes factor model comparison functions. These three models predicted our participants' intentionality judgments (nested within country) by the following:

1. **Belief model**: predict intentionality judgments by participant atheism (vs belief) alone, ignoring experimental condition.
    - $intent \sim 1 + atheism + (1 + atheism|country)$
2. **Experimental Condition model**: predict intentionality judgments by experimental condition (atheism vs belief) and participant atheism (vs belief).
    - $intent \sim 1 + condition + atheism + (1 + condition + atheism|country)$
3. **Interaction model**: predict intentionality judgments by experimental condition (atheism vs belief), participant atheism (vs belief), and their interaction.
    - $intent \sim 1 + condition + atheism + condition * atheism + (1 + condition + atheism + condition * atheism|country)$

Model comparisons allow direct inferences about evidence for or against each of our three key questions. All models predicted binary intentionality judgments from relevant predictors (condition and/or participant atheism), with all relevant intercepts and slopes being modeled as random across countries.

Do People Intuitively Prefer Belief to Atheism?

We first wanted to test whether people intuitively prefer religious belief to atheism overall – is there a Knobe effect for atheism? Comparing the Belief model to the Experimental Condition model allows a measure of evidence that experimental condition mattered. Indeed, there was strong evidence of a Knobe effect for atheism, $BF_{Condition,Belief} = 44.83$. This suggests quite strongly that experimental condition matters, but how?

Figure 1 displays posterior densities for the predicted probability of intentionality judgments across conditions and countries. Overall across countries, participants in the atheism condition were more likely to say that the newspaper story intentionally created atheists than participants in the religious belief condition were to say that the newspaper story intentionally created believers, $OddsRatio = 1.4$, [1.07, 1.85], $Pr(OR > 1) > .99$. This general pattern was evident in each individual country except Viet Nam, although it was less pronounced in China and Japan than the remaining 5 countries. Table 1 summarizes our findings, focusing on the coefficients and effect size indices for the random slopes of experimental condition across countries. By inference, this suggests that people intuitively view creation of atheists as worse (and thus more intentional) than creating believers. In short, these results are consistent with the existence of an intuitive preference for religious belief over atheism.

Table 1: Model results: Overall, and in 7 of 8 countries, model predicts higher intentionality judgments in the Atheist condition than the Theist condition. Note: Posterior modes (B and OR) are accompanied by highest posterior density intervals. Pr(B>0) is the posterior probability of the predicted directional difference.

Country	**B**	**B 0.89 HPDI**	**Odds Ratio**	**OR 0.89 HPDI**	**Pr(B > 0.00)**
Canada	0.24	[-0.01, 0.57]	1.24	[0.95, 1.70]	0.93
China	0.27	[-0.10, 0.49]	1.28	[0.88, 1.60]	0.87
Czechia	0.60	[0.23, 1.12]	1.76	[1.15, 2.80]	0.99
Japan	0.24	[-0.08, 0.51]	1.24	[0.86, 1.59]	0.86
Netherlands	0.57	[0.20, 0.96]	1.71	[1.16, 2.51]	0.99
Sweden	0.38	[0.06, 0.78]	1.41	[0.97, 2.06]	0.96
UK	0.72	[0.30, 1.24]	1.98	[1.32, 3.42]	> 0.99
Viet Nam	0.09	[-0.25, 0.35]	1.07	[0.74, 1.37]	0.58
Overall	**0.35**	**[0.10, 0.64]**	**1.40**	**[1.07, 1.85]**	**0.98**

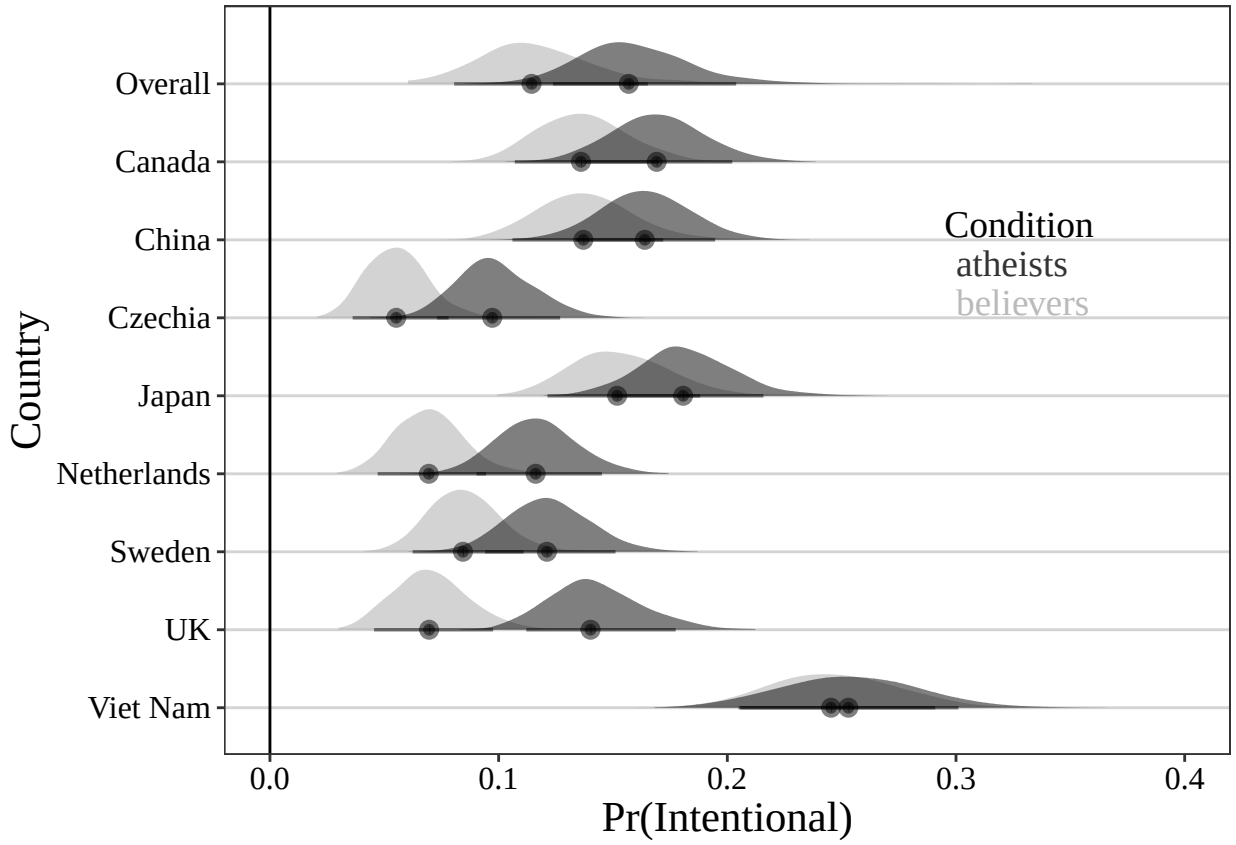

Figure 1: Posterior predictions for 'intentional' judgments, both overall and within each of the 8 countries. Everywhere except Viet Nam, people were more likely to rate the creation of atheists as more intentional, relative to the creation of theists. Note: height of curve indexes relative estimate credibility. Point estimates (posterior mode) and uncertainty (89% HPDI) also depicted for each estimate.

Do Intuitive Preferences for Religion Interact with Participant Atheism?

Initial analyses suggested a Knobe effect for atheism. Next, we sought to test whether this effect interacted with participant atheism. To evaluate the evidence for or against the possibility that the atheism Knobe effect was moderated by participant atheism, we compared the Interaction model to the Experimental Condition model. There was strong evidence against a possible interaction with participant atheism, $BF_{Condition,Interaction}$ = 10.38. The overall interaction slope was thoroughly consistent with zero, $\beta = 0.12$, [-0.24, 0.59]. This represents strong evidence against the possibility that the atheist Knobe effect meaningfully interacts with participant atheism.

Do Even Atheists Show an Intuitive Preference for Belief Over Atheism?

The first two sets of analyses revealed strong evidence (1) for an atheism Knobe effect, and (2) against an interaction with participant-level atheism. This strongly suggests a robust pattern whereby people – even atheists – intuitively prefer religious belief to atheism. To directly evaluate atheists' susceptibility to the atheism Knobe effect, we used the posterior from the interaction model and marginalized it to project the pattern of predicted results for participants who indicated that they don't believe in God (or gods). On average, even atheists thought that creating atheists was more intentional than creating believers, given our data and model. Table 2 shows posterior predicted effects, as marginalized to show us what we expect for atheist participants. Overall, our data and model suggest that even atheist participants across 8 secular countries tend to intuitively favor religion over atheism, $\beta = 0.49$, [0.1, 0.89], $Pr(\beta > 0)$ 0.97. This pattern was mirrored within each individual country (posterior probabilities ranging from .85 to .99).

Robustness Checks

We repeated this same series of analyses twice more, each using a different measure of participant religiosity and religious influence, to assess robustness of conclusions. Both sets of analyses yielded identical inferences, and are fully outlined in the Online Supplement. Across three measures of religiosity, there was again evidence *in favor* of an atheism Knobe effect, and evidence *against* an interaction with participant beliefs, consistent with an atheism Knobe effect even among participants scoring at floor on all three religiosity measures.

Discussion

We measured pro-religious intuitions – termed belief in belief by Dennett – in 8 largely secular countries with a variant of the Knobe effect (37) – an experimental philosophy finding which reveals that people rate negative

Table 2: Model results: Overall, and in all 8 countries, model predicts higher intentionality judgments among atheists in the Atheist condition than the Theist condition. Note: Posterior modes (B) are accompanied by highest posterior density intervals. Pr(B>0) is the posterior probability of the predicted directional difference.

Country	**B**	**0.89 HPDI**	**Pr(B > 0.00)**
Canada	0.42	[-0.15, 0.94]	0.86
China	0.35	[-0.27, 0.93]	0.85
Czechia	0.70	[0.07, 1.22]	0.96
Japan	0.41	[-0.17, 0.87]	0.87
Netherlands	0.53	[0.07, 1.12]	0.97
Sweden	0.42	[-0.16, 0.92]	0.87
UK	0.76	[0.25, 1.48]	0.99
Viet Nam	0.36	[-0.25, 1.19]	0.83
Overall	**0.49**	**[0.10, 0.89]**	**0.97**

side effects of actions as more intentionally caused than positive side effects. In our study, participants rated the religious shifts arising from a news article as more intentional if the article led to increased atheism rather than increased belief in God. This suggests that people intuitively view atheism more negatively than religious belief. There was strong evidence against this effect interacting with participant religiosity, and indeed the effect was evident even among atheist participants. Results were robust across analyses and alternative measures of participant beliefs. This atheism Knobe effect, largely general across participant atheism and across 8 largely secular countries, may suggest that belief in belief lingers, even in locales where explicit religious belief has considerably waned.

Why the Intuitive Self-Aversion?

These results are consistent with the possibility that even atheist participants in highly secular countries intuitively favor religious belief over atheism. Although perhaps surprising, this pattern of results is similar to those found in cross-cultural explorations of anti-atheist sentiments (20). In this previous project, even atheists in many secular countries were found to intuitively associate serial murder with atheism – a rather extreme form of intuitive distrust of one's fellow nonbelievers. What might explain intuitive aversion to atheism among atheists in secular countries?

We speculate that this is the result of a cultural evolutionary lag of sorts. Religions have exerted powerful influences for long stretches of cultural evolutionary time, and religious norms and institutions are immensely

potent even in countries that have recently secularized in an overt sense. People in these places may attend religious services less often, but they still see churches, synagogues, mosques, temples, and other places of worship all the time. People might no longer explicitly equate morality with religion, but cultural conflation of these concepts over time may have lingering effects on people's intuitions (9).

Some may speculate that the present results are consistent with an approach that views religious beliefs, and their deep connections to prosociality, as more innate than culturally evolved (47). However, this more nativist account would have trouble accommodating the cross-cultural varability in the present effects and similar previous ones (20). Further, models of religious cognition that do not interface closely with the full theoretical toolkit of cultural evolution struggle to account for other patterns in beliefs, practices, and differing relationships to morality and prosociality observed cross-culturally (10).

Relatedly, we have framed belief in belief as an intuition, and have also measured it using a task that captures intuitive preferences. Does this imply that belief in belief is innate, or in a sense unlearned? No. Some intuitions linked to religion (48–50) may come easily and naturally in development, and subsequently support the emergence of religious beliefs, given adequate cultural scaffolding and support (10). Other intuitive processes (like, we contend, belief in belief) emerge *as a result of* pre-existing religious influence – growing up in a world that has been heavily influenced by religious beliefs, practices, and institutions, cultural learners pick up on regular cultural associations between religious belief and moral goodness, and as a result tend to develop intuitive associations.

By situating belief in belief within a broader theoretical framework for understanding how norms, beliefs, institutions, and intuitions interact and change across and within cultures, this account suggests some exciting avenues for future research. To the extent that latent effects of religion persist at an intuitive level, then it should be possible to find intuitive traces in other domains typically associated with and influenced by religions. Here we found evidence of a general preference; in prior work we found evidence in the domain of moral intuitions (20); others have speculated that patterns in work and childrearing in the USA reflect an intuitive remnant of overt Protestantism (51). These domains and others merit focused empirical attention.

Caveats, Alternatives, and Limitations

We found evidence consistent with the possibility that even atheists in highly nonreligious societies nonetheless intuitively prefer belief to atheism, an intuition captured by the Knobe effect paradigm. However, we should acknowledge that the Knobe effect task is a bit of an odd one, potentially amenable to multiple explanations. Here we consider two possible alternative explanations for our findings, without necessarily endorsing the possibilities outlined.

First, in addition to our predicted dynamics concerning lingering pro-religious intuitions, our results are also compatible with the possibility that this task does not measure subtle intuitions about religious belief, per se, but instead reflects an as-yet unidentified methodological asymmetry in how our vignettes were received. For example, participants might be imagining wholly different scenarios when contemplating a news story that could result in people gaining or losing religious faith – perhaps, for example, people inferred that a news story leading to widespread atheism would necessarily be more negative somehow than a story leading faith to flourish. This difference in imagined vignette contents, rather than intuitions about religion, may have caused a difference in intentionality judgments. However, as we see it, this skeptical interpretation can be leveled against most social science and experimental philosophy vignette studies (including previous work on the Knobe effect itself), and is not a unique shortfall of our study. Nonetheless, our proposed explanation – that the Knobe effect tracks intuitions about the relative positivity and negativity that people hold towards atheism and belief, respectively – is supported by other theory and evidence. We merely entertain the possibility that there is more (or less) to the Knobe effect task than initially assumed. That said, if there is a widespread intuition that events promoting atheism must themselves be negative, that seems to fit the broader dynamics that underpin our hypotheses – an intuitive negative association with atheism, even among atheists. It simply transfers the negativity from the creation of atheists to the (presumably horrific) events that could spawn them.

Second, another plausible alternative explanation involves normative considerations. As discussed within the broader experimental philosophy literature (52–54), behavior that violates norms (moral or otherwise) is more informative about underlying mental states than norm-conforming behavior is. To adapt Uttich and Lombrozo's neat illustration (52), we can infer more about a person's preference for academic regalia if they wear it at the beach (which would be norm-violating) than if we notice them wearing it at a graduation ceremony (which would be norm-conforming). In this view, participants might attribute more intentionality to actions leading to atheism not because they intuitively favor religious belief, but because they perceive such actions as violating societal norms favoring religious belief. We believe that while this alternative explanation is worth considering, the robustness of the Knobe effect we observed across diverse cultural contexts – even in highly nonreligious societies, where pro-religious norms are presumably weak – supports our original interpretation that individual preferences play a significant role.

The present Knobe effect results are consistent with the possibility that even atheists in nonreligious societies intuitively disfavor atheism, relative to belief – as well as with other possible explanations. Similarly, our prior work using the representativeness heuristic is consistent with the possibility that there is substantial intuitive distrust of atheists, even among atheists in nonreligious societies (20) – as well as with other possible explanations. Both tasks – with their acknowledged idiosyncracies – triangulate towards similar inferences

about the persistence of pro-religious intuitions, even in the face of declines in explicit and overt religiosity. This evidentiary convergence across tasks points to further possibilities for studying intuitive preferences for belief and against atheism beyond the spheres in which we've thus far investigated.

We sought to explore pro-religious intuitions in a diverse set of largely secular countries, but emphasize that our sampling precludes strong inference about these sorts of intuitions in other countries. By sampling countries that varied in majority historical religious background, we hoped to allay concerns that any observed effects were quirks of particular religious traditions. However, we could not fully dismiss this possibility; and questions of which religions are more prone to secularization, and the possibility that some might be more susceptible to the lingering of latent pro-religious intuitions, are well worth pursuing in future research.

Coda

Chunks of the world are rapidly secularizing, at least according to measures of religious attendance, prayer frequency, and self-reported belief in gods (2, 3). But religions have enjoyed prominence for most of our recent cultural evolutionary history, and have powerfully shaped our norms and institutions (12, 36). This may have created a sort of cultural lag, whereby some beliefs and intuitions shift more quickly than others over cultural evolutionary time. We found evidence that rapid overt secularization might overlay latent pro-religious intuitions – a sort of lingering intuitive cultural influence of religion. Although atheists in secular countries have abandoned belief, they appear to have retained a modicum of intuitive belief in belief.

Materials and Methods

Sample

In total, we recruited a sample of 5400 people from 8 largely secular countries: Canada, China, Czechia, Japan, the Netherlands, Sweden, the UK, and Viet Nam. Sampling used quotas to approximate representativeness on key demographics, but sample size and use of quotas preclude strong claims thereof. A total of 3804 participants remained after we dropped participants who had failed at least one of two included attention checks. The final sample was 51% men, with an average age of 48. Per-country sample sizes were comparable, ranging from 436 to 502. Additional sampling information is available in the Supplement.

Procedure

We measured intuitive preferences for religion (over atheism) with a Knobe effect vignette, and also religious demographics.

Intuitive Preferences

We constructed two different vignettes to infer belief in belief, by assessing the degree to which demographic shifts towards faith and atheism, respectively, are seen as intentionally caused. Here is the atheist condition vignette:

> *A journalist went to her editor and said "I have written a news story about religion. I'm sure that it will help us sell newspapers and make money, and I think it will also make some people believe that God doesn't exist." The editor answered, "I don't care at all if people believe that God doesn't exist. I just want to make as much money as I can. Let's publish this news story." The newspaper published the news story. After reading the news story some people who did believe in God changed their mind and started to believe that God doesn't exist.*
>
> **Did the editor intend to make people believe that God doesn't exist?**
>
> **YES** or **NO**

For completeness, the religious believer condition read:

> *A journalist went to her editor and said "I have written a news story about religion. I'm sure that it will help us sell newspapers and make money, and I think it will also make some people believe that God exists." The editor answered, "I don't care at all if people believe that God exists. I just want to make as much money as I can. Let's publish this news story." The newspaper published the news story. After reading the news story some people who didn't believe in God changed their mind and started to believe that God exists.*
>
> **Did the editor intend to make people believe that God exists?**
>
> **YES** or **NO**

Religious Demographics

In addition to this focal measure of intuitive preferences for religious belief over atheism, we also assessed participants' own religious beliefs. We were especially interested in the question of whether even atheists might show an intuitive preference for religious belief. One demographic item asked participants about their belief in God, providing the following response options:

- I believe in God (or gods)
- I don't know whether or not God (or gods) exists
- I don't really take a strong stance on God (or gods)

- I don't believe in God (or gods)

Because our primary hypothesis related to atheism, we coded this item as a binary, pitting staunch atheists ("I don't believe in God (or gods)") against all other response options. For robustness, we ran parallel analyses using two additional measures related to participant religiosity. One was the the 6-item version of the Supernatural Belief Scale (55, 56), and the other assessed how much participants had witnessed credibility enhancing displays (16) of their parent's religious faith while growing up (27). In previous research, this latter item proves to be an excellent predictor of adult atheism (30, 31, 57–59). In our sample, both scales proved fairly reliable, $\omega_h = .91$ and $\omega_h = .89$, respectively. We thus were able to test whether even fairly directly measured atheists (using the binary measure) also show intuitive favoritism for religious belief, and additionally test robustness with two conceptually related operationalizations of religious belief and cultural exposure to religion.

Transparency Statement

In the spirit of full transparency: this project was one portion of a larger data collection effort, testing a number of distinct hypotheses. Specifically, a large cross-cultural survey was fielded, primarily to look at cognitive and cultural correlates of nonbelief in eight largely secular countries. When there was spare room in the survey a team member suggested adding our focal Knobe effect item as a clearly distinct hypothesis to test, entirely separate from the other measures in the study. Our first stage of data processing for this paper involved creating a data frame that only contained our focal variable and related demographics. To date, the larger survey has not had its variables coded and compiled, and no analyses have been done on the main dataset.

Data and analysis code are available at https://osf.io/gxft8/?view_only=4bba4f6ad7d24d6685943be6096f0b11

Author Note: This paper arose as part of a larger collaboration among all authors. WMG and RTM wrote the initial draft, based on analyses by WMG. All authors contributed to study design and paper revision. We acknowledge the following financial support: John Templeton Foundation (#60624; #61928, to all; #62631, to RMR), the NOMIS Foundation ("Collective Delusions: Social Identity and Scientific Misbeliefs", to RTM), and a Leverhulme International Professorship Grant (LIP-2022-001, RTM).

References

1. D. C. Dennett, *Breaking the spell: Religion as a natural phenomenon* (Penguin, 2006).

2. P. Norris, R. Inglehart, *Sacred and secular: Religion and politics worldwide* (Cambridge University Press, 2011).

3. R. F. Inglehart, *Religion's sudden decline: What's causing it, and what comes next?* (Oxford University Press, 2020).

4. A. Norenzayan, *Big gods: How religion transformed cooperation and conflict* (Princeton University Press, 2013).

5. A. Norenzayan, *et al.*, The cultural evolution of prosocial religions. *Behavioral and Brain Sciences* **39**, 1–65 (2016).

6. S. Atran, J. Henrich, The Evolution of Religion: How Cognitive By-Products, Adaptive Learning Heuristics, Ritual Displays, and Group Competition Generate Deep Commitments to Prosocial Religions. *Biology Theory* **5**, 18–30 (2010).

7. B. G. Purzycki, *et al.*, Moralistic gods, supernatural punishment and the expansion of human sociality. *Nature* (2016).

8. Q. D. Atkinson, P. Bourrat, Beliefs about God, the afterlife and morality support the role of supernatural policing in human cooperation. *Evolution and Human Behavior* **32**, 41–49 (2011).

9. R. McKay, H. Whitehouse, Religion and morality. *Psychological bulletin* **141**, 447 (2015).

10. W. M. Gervais, *Disbelief: The Origins of Atheism in a Religious Species* (Rowman & Littlefield, 2024).

11. J. Henrich, Cultural group selection, coevolutionary processes and large-scale cooperation. *Journal of Economic Behavior & Organization* **53**, 3–35 (2004).

12. J. Henrich, *et al.*, Markets, Religion, Community Size, and the Evolution of Fairness and Punishment. *Science* **327**, 1480–1484 (2010).

13. J. Henrich, *et al.*, Costly punishment across human societies. *Science* **312**, 1767–1770 (2006).

14. N. Baumard, P. Boyer, Explaining moral religions. *Trends in Cognitive Sciences* **17**, 272–280 (2013).

15. S. S. Wiltermuth, C. Heath, Synchrony and cooperation. *Psychological Science* **20**, 1–5 (2009).

16. J. Henrich, The evolution of costly displays, cooperation and religion. *Evolution and Human Behavior* **30**, 244–260 (2009).

17. D. Xygalatas, *et al.*, Extreme rituals promote prosociality. *Psychological science* **24**, 1602–1605 (2013).

18. C. Tamir, A. Connaughton, A. M. Salazar, The global God divide. *Pew Research Center* **20** (2020).

19. J. L. Brown-Iannuzzi, S. McKee, W. M. Gervais, Atheist horns and religious halos: Mental representations of atheists and theists. *Journal of Experimental Psychology: General* **147**, 292 (2018).

20. W. M. Gervais, *et al.*, Global evidence of extreme intuitive moral prejudice against atheists. *Nature Human Behaviour* **1**, s41562-017-0151 (2017).

21. W. M. Gervais, Everything is permitted? People intuitively judge immorality as representative of atheists. *PLoS One* **9**, e92302 (2014).

22. A. Dayer, *et al.*, Intuitive moral bias favors the religiously faithful. *Scientific reports* **14**, 18291 (2024).

23. A. Norenzayan, W. M. Gervais, The origins of religious disbelief. *Trends in cognitive sciences* **17**, 20–25 (2013).

24. D. Voas, The rise and fall of fuzzy fidelity in Europe. *European Sociological Review* **25**, 155–168 (2009).

25. D. C. Dennett, L. LaScola, *Caught in the pulpit: Leaving belief behind* (Pitchstone Publishing (US&CA), 2015).

26. W. M. Gervais, M. B. Najle, Learned faith: The influences of evolved cultural learning mechanisms on belief in Gods. *Psychology of Religion and Spirituality* **7**, 327 (2015).

27. J. Lanman, M. D. Buhrmester, Religious actions speak louder than words: Exposure to credibility-enhancing displays predicts theism. *Religion, Brain & Behavior* **7**, 3–16 (2017).

28. W. M. Gervais, A. K. Willard, A. Norenzayan, J. Henrich, The cultural transmission of faith: Why innate intuitions are necessary, but insufficient, to explain religious belief. *Religion* **41**, 389–410 (2011).

29. J. Lanman, The importance of religious displays for belief acquisition and secularization. *Journal of Contemporary Religion* **27**, 49–65 (2012).

30. A. K. Willard, L. Cingl, Testing theories of secularization and religious belief in the Czech Republic and Slovakia. *Evolution and Human Behavior* **38**, 604–615 (2017).

31. W. M. Gervais, M. Najle, N. Caluori, The Origins of Religious Disbelief: A Dual Inheritance Approach. *Social Psychological and Personality Science* (2020).

32. H. D. Turpin, *Unholy Catholic Ireland: Religious hypocrisy, secular morality, and Irish irreligion* (Stanford University Press, 2022).

33. Pew, Religious Landscape Study. *Pew Research Center's Religion & Public Life Project* (2014) (January 10, 2024).

34. P. J. Richerson, R. Boyd, *Not by genes alone: How culture transformed human evolution* (University of Chicago press, 2008).

35. J. Henrich, *The WEIRDest people in the world: How the West became psychologically peculiar and particularly prosperous* (Penguin UK, 2020).

36. J. F. Schulz, D. Bahrami-Rad, J. P. Beauchamp, J. Henrich, The Church, intensive kinship, and global psychological variation. *Science* **366** (2019).

37. J. Knobe, Intentional action and side effects in ordinary language. *Analysis* **63**, 190–194 (2003).

38. A. M. Leslie, J. Knobe, A. Cohen, Acting intentionally and the side-effect effect - Theory of mind and moral judgment. *Psychological Science* **17**, 421–427 (2006).

39. J. R. Beebe, W. Buckwalter, The epistemic side-effect effect. *Mind & Language* **25**, 474–498 (2010).

40. A. Gelman, J. Hill, M. Yajima, Why we (usually) don't have to worry about multiple comparisons. *Journal of Research on Educational Effectiveness* **5**, 189–211 (2012).

41. R. McElreath, *Statistical Rethinking: A Bayesian Course with Examples in R and Stan* (CRC Press, 2016).

42. R. McElreath, Multilevel Regression as Default | Elements of Evolutionary Anthropology (2017) (January 16, 2020).

43. R. Hoekstra, R. D. Morey, J. N. Rouder, E.-J. Wagenmakers, Robust misinterpretation of confidence intervals. *Psychonomic Bulletin & Review* **21**, 1157–1164 (2014).

44. R. D. Morey, R. Hoekstra, J. N. Rouder, M. D. Lee, E.-J. Wagenmakers, The fallacy of placing confidence in confidence intervals. *Psychonomic bulletin & review* **23**, 103–123 (2016).

45. J. K. Kruschke, Bayesian analysis reporting guidelines. *Nature human behaviour* **5**, 1282–1291 (2021).

46. P.-C. Bürkner, Brms: An R package for Bayesian multilevel models using Stan. *Journal of statistical software* **80**, 1–28 (2017).

47. D. Johnson, *God is watching you: How the fear of God makes us human* (Oxford University Press, 2015).

48. D. Kelemen, Are children "intuitive theists"? Reasoning about purpose and design in nature. *Psychological Science* **15**, 295–301 (2004).

49. D. Kelemen, E. Rosset, The human function compunction: Teleological explanation in adults. *Cognition* **111**, 138–143 (2009).

50. P. Bloom, Religion is natural. *Developmental science* **10**, 147–151 (2007).

51. E. L. Uhlmann, T. A. Poehlman, J. A. Bargh, Implicit theism. *Handbook of motivation and cognition across cultures*, 71–94 (2008).

52. K. Uttich, T. Lombrozo, Norms inform mental state ascriptions: A rational explanation for the side-effect effect. *Cognition* **116**, 87–100 (2010).

53. F. Hindriks, Normativity in action: How to explain the Knobe effect and its relatives. *Mind & Language* **29**, 51–72 (2014).

54. J. Knobe, Reason explanation in folk psychology. *Midwest Studies in Philosophy* **31**, 90–106 (2007).

55. J. Jong, M. Bluemke, J. Halberstadt, Fear of death and supernatural beliefs: Developing a new Supernatural Belief Scale to test the relationship. *European Journal of Personality* **27**, 495–506 (2013).

56. J. Jong, *Death anxiety and religious belief* (Bloomsbury Publishing, 2016).

57. David. L. R. Maij, *et al.*, Mentalizing skills do not differentiate believers from non-believers, but credibility enhancing displays do. *PLoS ONE* **12**, e0182764 (2017).

58. H. D. Turpin, A. K. Willard, Credibility Enhancing Displays, religious scandal and the decline of Irish Catholic orthodoxy. *Evolutionary Human Sciences* **4**, e20 (2022).

59. H. D. Turpin, M. Andersen, J. A. Lanman, CREDs, CRUDs, and Catholic scandals: Experimentally examining the effects of religious paragon behavior on co-religionist belief. *Religion, Brain & Behavior* **9**, 143–155 (2019).