REGISTERED REPORT STAGE 2

the british
psychological society
promoting excellence in psychology

'So Help Me God'? Does oath swearing in courtroom scenarios impact trial outcomes?

Ryan T. McKay[1] | Will Gervais[2] | Colin J. Davis[3]

[1]Department of Psychology, Royal Holloway, University of London, Egham, UK

[2]Centre for Culture and Evolution, Brunel University London, Uxbridge, UK

[3]School of Psychological Science, University of Bristol, Bristol, UK

Correspondence

Ryan T. McKay, Department of Psychology, Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK.
Email: ryan.mckay@rhul.ac.uk

Abstract

In countries such as Britain and the US, court witnesses must declare they will provide truthful evidence and are often compelled to publicly choose between religious ("oath") and secular ("affirmation") versions of this declaration. Might defendants who opt to swear an oath enjoy more favourable outcomes than those who choose to affirm? Two preliminary, pre-registered survey studies using minimal vignettes (Study 1, $N = 443$; Study 2, $N = 913$) indicated that people associate choice of the oath with credible testimony; and that participants, especially religious participants, discriminate against defendants who affirm. In a third, *Registered Report* study (Study 3, $N = 1821$), we used a more elaborate audiovisual mock trial paradigm to better estimate the real-world influence of declaration choice. Participants were asked to render a verdict for a defendant who either swore or affirmed, and were themselves required to swear or affirm that they would try the defendant in good faith. Overall, the defendant was not considered guiltier when affirming rather than swearing, nor did mock-juror belief in God moderate this effect. However, jurors who themselves swore an oath *did* discriminate against the affirming defendant. Exploratory analyses suggest this effect may be driven by authoritarianism, perhaps because high-authoritarian jurors consider the oath the traditional (and therefore correct) declaration to choose. We discuss the real-world implications of these findings and conclude the religious oath is an antiquated legal ritual that needs reform.

KEYWORDS

atheism, legal processes, morality, prejudice, religious beliefs

BACKGROUND

It is simply impossible for people to be moral without religion or God.
~ Laura Schlessinger (quoted in Zuckerman, 2008, p. 6)

Suppose a witness to be the worst of Infidels… If he is honest enough to subject himself to the disability, rather than tell a lie, why exclude him?
~ Justice Scott, writing for the Supreme Court of Virginia (Grattan, 1847, p. 642)

"I swear by Almighty God that the evidence I shall give shall be the truth, the whole truth and nothing but the truth." With this familiar phrase, recited daily in some form in courtrooms throughout the world, trial witnesses invoke a supernatural power to strengthen the credibility of their evidence. In countries such as Britain, Australia and the United States, however, witnesses can opt for a secular version of this declaration. Rather than citing God as their witness, they can instead "solemnly, sincerely and truly declare and affirm" that they will tell the truth. From a legal perspective, both versions, religious ("oath") and secular ("affirmation"), are equally binding. But is the secular affirmation as effective as the religious oath in conveying trustworthiness to judge and jury?

Moral suspicion of atheists is globally widespread and deeply entrenched. In Britain, 20% of survey respondents explicitly agree with Laura Schlessinger that morality is impossible without belief in God, while this attitude is far more prevalent in the US (44%) and more prevalent still in many other countries (Pew Research Center, 2020). In a cross-national study spanning 13 diverse nations, Gervais et al. (2017) confirmed that distrust of atheists is pervasive and intuitive even for non-believers. Participants in most of these countries, including Britain and the US, were more likely – roughly twice as likely overall – to view immoral behaviour as representative of atheists, relative to religious believers. This moral prejudice against non-believers was evident even in those who professed complete disbelief in God (see also Edgell et al., 2006; Gervais, 2011, 2013, 2014; Gervais et al., 2011; Gervais & Norenzayan, 2012, 2013; Giddings & Dunn, 2016; Hughes et al., 2015; Tan & Vogel, 2008).

Moral prejudice against atheists has important implications in the legal system, as it carries the potential to bias juridical decisions. Although legal formalists may insist that such decisions are made with dispassionate deliberation (Danziger et al., 2011), a range of studies suggest they can be influenced by individual characteristics, experiences and other immaterial factors (Englich et al., 2006; Kang et al., 2012; Simon, 2012; Yamamoto et al., 2019). Glynn and Sen (2015), for instance, document a liberalizing "daughters effect" among US Courts of Appeals judges voting in gender-related cases: Conditional on the number of children a judge has, male judges with daughters vote in a more feminist-leaning fashion on gender issues than those who only have sons (see also Boyd et al., 2010). Legal decisions are also known to be compromised by racial biases (Eren & Mocan, 2018; Hunt, 2015; Mitchell et al., 2005; Sommers & Marotta, 2014). Also, Cho et al. (2017) even found that sentences rendered in US federal courts on "sleepy Monday" (the first Monday after the spring shift to daylight saving time in the US) were more punitive than those dispensed on comparison Mondays (but see Spamann, 2018, for a critique of this study).

As for religion, in their book *God in the Courtroom*, Bornstein and Miller (2009) concluded that while religious factors are less important than the facts of a given case in determining trial outcomes, there are nevertheless numerous cases in which such factors can be decisive. One of their own studies showed that mock jurors were least punitive when a defendant was described as having converted to Christianity, compared to when the defence attorney made a generic appeal for Christian forgiveness (Miller & Bornstein, 2006). Subsequent mock jury studies have documented how jurors' religious characteristics influence their verdicts and sentencing decisions (Miller et al., 2014). And in recent work, Brown-Iannuzzi et al. (2021) found that participants perceived a Christian rape victim as more moral than an atheist victim, which predicted a higher conviction rate.

This brings us back to the possible effects of a witness's declaration choice on juridical observers. In legal systems where witnesses choose between religious and secular declarations when being sworn in, this choice may represent judges' and jurors' first impression of a witness. The upshot is that when witnesses are called to the stand – where perceived credibility is paramount – they may be compelled by legal procedure to signal their belief or disbelief in God. Given entrenched distrust of non-believers, the risk is that witnesses who opt for the affirmation may appear less credible than those who choose the oath, biasing trial outcomes in any number of ways. Most obviously, defendants who choose the oath when giving evidence may enjoy more favourable verdicts and sentencing decisions than those who opt for the affirmation. The purpose of the present research was to investigate the potential for such bias.

An important caveat to note at this point, however, is that in many cases, witnesses may opt for the affirmation despite sincere belief in God; indeed, some witnesses (e.g. Quakers, Mennonites) may choose to affirm *for religious reasons*, perhaps because of adherence to scriptural passages interpreted as prohibiting oaths (e.g. *Matthew 5:34–37, James 5:12*). The historical origins of the affirmation lie in the refusal of Quakers to swear oaths: In 1696, following the English Parliament's Quakers Act 1695, Quakers were permitted to make an affirmation instead of swearing an oath (Maitland, 1908/2008). It is possible, therefore, that jurors may not see the affirmation as a signal of disbelief in God; on the contrary, this choice could be viewed as a signal of religious integrity.[1] Accordingly, in our first study, we sought evidence of the connection – and crucially, the *perceived* connection – between declaration choice and religious belief or lack thereof.

STUDY ONE

Research questions

The primary hypothesis of this study was that witnesses described as choosing to swear an oath would be perceived as more religious than those described as choosing to make an affirmation. In addition, we sought to explore participants' stated reasons for their own declaration choices, and to investigate whether those who choose the oath in court are actually more religious than those who choose to affirm.

Method

Participants

We collected data from 443 participants via the online platform *Prolific* (https://www.prolific.co/). Our survey was produced using *Qualtrics* (https://www.qualtrics.com). All participants were British citizens and residents (75% female, Mean age = 37.5 years [$SD = 11.8$]). We required 382 participants to attain 90% power ($a = .05$) to detect a small effect ($d = 0.15$) in our primary analysis (one-tailed related samples t-test). We added approximately 15% to this number to account for planned participant exclusions (any participants who skipped survey questions; none did). The hypotheses, design, data collection and analysis plan for our first two studies were preregistered with *AsPredicted* (Study 1: https://aspredicted.org/iy2hv.pdf[2]; Study 2: https://aspredicted.org/ra6m5.pdf). De-identified data files and analysis scripts for both studies are available on the Open Science Framework: https://osf.io/rk8ds/?view_only=None. Ethical approval was obtained through the self-certification process at Royal Holloway, University of London.

---

[1]Indeed, this is an argument that we have faced repeatedly when suggesting that the oath/affirmation distinction could potentially bias courtroom decisions – see McKay and Davis (2017) and associated reader comments (e.g. "affirming was introduced to meet the principled objections of Quakers and a few fundamentalist Christians who adhered to Jesus' instruction not to swear oaths… It does not distinguish between Christians and atheists").

[2]Due to an error by the first author, the pre-registration for Study 1 erroneously indicates that some data had already been collected at the point of pre-registration. This was not in fact the case.

Procedure

Participants were asked whether they had ever given evidence in a British court. They were then informed/reminded of the declaration options available to witnesses and defendants in British courts (we presented the exact text of the oath and affirmation), and asked which declaration they had chosen ($N = 70$[3]), or would hypothetically choose ($N = 373$). In each case, there were three options: "oath", "affirmation" and "don't know". We also asked them to explain the reason(s) for their choice.

We then presented participants with brief information about two hypothetical court witnesses, "Sam" and "Pat". One witness was described as choosing to take an oath, the other as choosing to make an affirmation (we counterbalanced whether it was Sam or Pat who took the oath, and also counterbalanced the order in which the oath-taking witness and affirmation-taking witness were presented; the texts of the oath and affirmation remained onscreen for reference). In each case, participants were asked to indicate the witness's probable level of religious belief using a slider anchored by 1 ("Strongly Atheist") and 5 ("Strongly Religious"). Finally, participants indicated their age and gender, as well as their own religious affiliation[4] and belief in God (the latter on a 0–100 scale).

Results

Analyses for our studies were carried out in R (R Core Team, 2021). Hypothetical witnesses described as choosing the oath were perceived as much more religious (mean perceived level of religious belief = 3.92, $SD = 0.64$) than those described as choosing the affirmation (mean perceived belief = 2.34, $SD = 0.78$), $t(442) = 29.9$, $p < .001$, $d = 2.22$. There is some evidence that this effect was moderated by perceiver religiosity: although affirmation-choosing witnesses were viewed as much less religious than oath-choosing witnesses by affiliated and unaffiliated participants alike, this was especially the case for unaffiliated participants, $F(1, 398) = 12.76$, $p < .001$, $\eta_G^2 = .02$.[5]

The perception that declaration choice reflects religiosity seems generally accurate, as participants who themselves chose the oath were much more religious (Mean belief in God = 55.55, $SD = 36$, $n = 150$) than those who chose the affirmation (Mean belief = 18.45, $SD = 27.63$, $n = 275$), $t(246.34) = 10.98$, $p < .001$, $d = 1.2$.[6] There was also a strong association between the participants' chosen declaration and whether or not they had a religious affiliation, $\chi^2(1) = 135.64$, $p < .001$ (see Table 1). The odds of a person having a

TABLE 1 Contingency table showing Study 1 participants' chosen declarations as a function of their religious affiliation or lack thereof.

	Participant's affiliation status		
	Affiliated	Unaffiliated	Total
Participant's chosen declaration			
Oath	115	27	142
Affirmation	48	192	240
Total	163	219	382

---

[3]Participants who had given evidence on more than one occasion were instructed to think of the most recent occasion.

[4]We used these responses to compute a binary affiliation variable, with participants who indicated a Christian, Buddhist, Muslim, Hindu or Jewish affiliation coded as religiously affiliated ($n = 174$), and those who indicated "None" or "Atheist" coded as unaffiliated ($n = 226$). Those indicating "Agnostic" or "Other" for the affiliation question were excluded for purposes of analyses with this binary variable.

[5]However, the strength of perceivers' beliefs in God did not significantly moderate the relationship between their perceptions of oath-choosers and their perceptions of affirmation-choosers ($p = .169$).

[6]This was the case both for participants who had actually made this decision in court (Oath: Mean belief = 48.08, $SD = 39.57$, $n = 37$; Affirmation: Mean belief = 12.58, $SD = 34.59$, $n = 24$; $t(57.47) = 4.53$, $p < .001$, $d = 1.05$) and for those who answered hypothetically (Oath: Mean belief = 58, $SD = 34.59$, $n = 113$; Affirmation: Mean belief = 19.01, $SD = 28.13$, $n = 251$; $t[181.43] = 10.52$, $p < .001$, $d = 1.29$). Note that participants who answered "do not know" for the declaration question were excluded for purposes of these analyses.

religious affiliation were 16.87 (95% CI = [9.79, 29.89]) times higher if they chose the oath than if they chose the affirmation.[7]

These findings provide evidence that the declaration a witness chooses in court is a clear signal – and, crucially, is *perceived* as a clear signal – of religiosity or lack thereof.

Reasons for choice

We used a mixed methods approach to explore participants' stated reasons for their declaration choices (for full details see the Supporting Information). To summarize, we first examined the reasons provided and identified a set of emergent themes. We then recruited two graduate students (blind to the study aims and hypotheses) to independently code the responses into these designated thematic categories. The initial agreement between the independent coders was substantial ($\varkappa$ = .66): they subsequently met to resolve any discrepancies and to produce final agreed categorizations.

Consistent with the above-reported association between participants' chosen declaration and religious affiliation/belief, the most common reason for choosing a declaration was religious affiliation/belief or lack thereof (for choice of the oath and affirmation, respectively; no participants stated that they chose the affirmation because of their religion). However, participants who chose the oath were much less likely to cite their religious belief (36%) than participants who chose the affirmation were to cite their *lack* of belief (85%). For those who chose the oath, a full 20% did so because they believed it was the more credible choice (e.g. "I believed [the oath] would be perceived as more credible by the jury"). In contrast, just 4% of those who chose the affirmation did so for reasons of credibility (e.g. "[the affirmation is] more likely to be believed as not everyone believes in God"). Interestingly, one person who chose the affirmation noted that they "might swear an oath if I was planning on lying". Participants also chose the oath because it was more familiar to them (12%), because they saw it as the more "traditional" alternative (15%), or because they simply had not been given a choice in court (11%). Finally, 10% of participants who chose the affirmation did so because they believed the oath was inappropriate (e.g. "religion should have no place in politics or law"); although, as above, none explicitly stated a *religious* opposition to the oath.

These findings indicate that people associate choice of the oath with credible testimony. In Study 2, we sought to test this perception more directly using an experimental paradigm.

STUDY TWO

Research questions

The purpose of our second study was to investigate whether the type of legal declaration made by defendants in a trial – either religious (oath) or secular (affirmation) – can influence perceptions of their probable guilt. We pre-registered two key hypotheses:

H1: A defendant described as choosing to make an affirmation would be perceived as more likely to be guilty than if described as choosing to swear an oath.
H2: The effect in H1 would be moderated by perceiver religiosity, such that religious onlookers would be more likely to associate the affirmation with guilt than would non-religious onlookers.

---

[7]The association between declaration and affiliation was strong whether participants had actually made this decision in court, $\chi^2(1) = 26.33$, $p < .001$; Odds ratio = 25.51 [5.62, 153.12] or had answered hypothetically, $\chi^2(1) = 104.1$, $p < .001$; Odds ratio = 14.93 [8.22, 28.09].

Method

Participants

We collected data from 915 participants, again via *Prolific*. We required 788 participants to attain 80% power ($a = .05$) to detect a small effect ($d = 0.2$) in our primary analysis (two-tailed independent samples *t*-test or non-parametric alternative). We added approximately 15% to this number to account for planned participant exclusions (any participants who skipped survey questions; $N = 2$), leaving a final $N = 913$ (71% female, Mean age $= 36.1$ years [$SD = 11.7$]). All participants were British citizens and residents.

Procedure

Participants were directed to a *Qualtrics* survey where they were presented with a brief description of the declaration options available to defendants in British trials (we presented the exact text of the oath and affirmation and counterbalanced whether the oath or affirmation was described first, lest participants associate probable guilt with not taking the first option described). We then presented brief information about a defendant in a murder trial, and asked participants to indicate his probable guilt. Participants were randomly allocated to one of two between-subjects declaration conditions (oath or affirmation) as follows:

> Alan (not his real name) is a 37-year-old British man who recently stood trial for the murder of his wife. Assuming Alan chose to [swear an oath/make an affirmation], how likely do you think it is that he is guilty of the murder of his wife? Please use the following 0–100 scale to indicate Alan's probable level of guilt.

In both conditions participants used a slider, anchored by *Very likely innocent* and *Very likely guilty*, to indicate the defendant's probable guilt. The slider ranged from 0–100 and the exact response was displayed as participants moved it, so that participants wishing to select the exact midpoint of the scale could easily do so.

Finally, participants indicated their age and gender, as well as their own religious affiliation and belief in God (the latter on a 0–100 scale).

Results

Five hundred and nineteen participants (57% of the sample) sat precisely on the fence, giving a probable guilt rating of 50 (see Figure 1). As this pronounced spike in the distribution violates parametric assumptions (neither the variable nor the residuals are normally distributed), we report non-parametric bootstrapped confidence intervals alongside parametric statistics below.

As predicted, the defendant was perceived as slightly more likely to be guilty when described as choosing to affirm (mean guilt rating $= 54.28$, $SD = 14.2$, $n = 456$) than when described as choosing to swear an oath (mean $= 52.19$, $SD = 13.72$, $n = 457$), $t(909.8) = 2.26$, $p = .024$, $d = 0.15$; bootstrapped *t*-test with 10,000 resamples, BCa 95% CI for mean difference: [0.33, 3.92]; BCa 95% CI for d estimate: [0.02, 0.28].

This effect was moderated by perceiver religiosity: religious believers were more affected by the defendant's choice of declaration than non-believers were (see Table 2 and Figure 2). We computed simple slopes to examine the effect of the observer's belief in God on the perceived guilt of the defendant in each declaration condition. There was no significant effect for defendants who chose to swear an oath (bootstrapped regression with 10,000 resamples, BCa 95% CI: [−0.06, 0.02]), but the higher the observer's belief in God, the guiltier they perceived defendants who chose to make an affirmation to be (bootstrapped 95% CI: [0.05, 0.13]).

This pattern was replicated in the religious affiliation moderation analysis (see Table 3 and Figure 3). To investigate whether having a religious affiliation moderated the effect of declaration condition on guilt perceptions, we computed a binary affiliation variable in the same way as for Study 1 (see fn. 4). We examined simple effects of declaration condition on perceptions of probable guilt at each level of this variable.
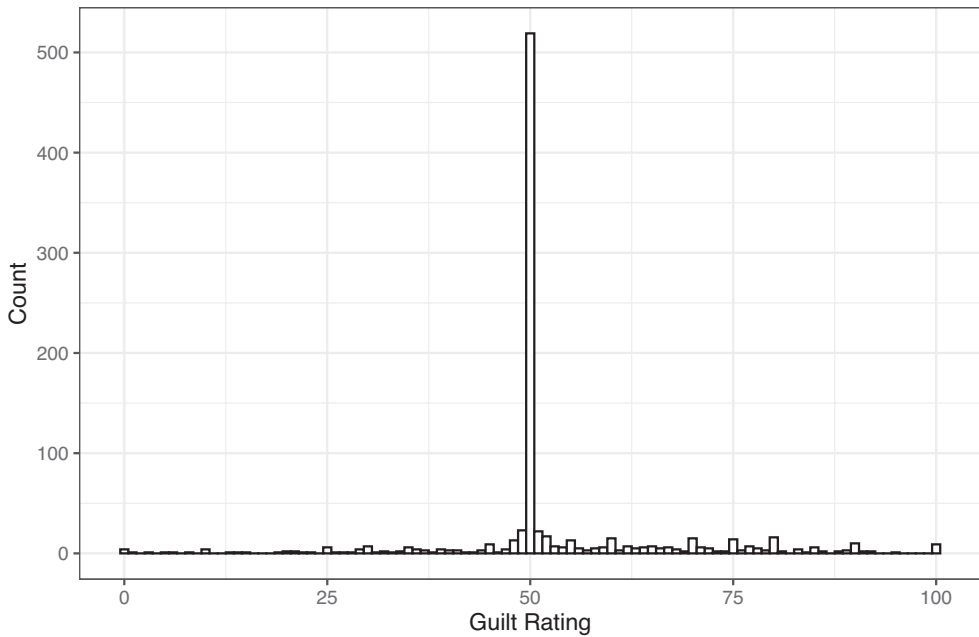
FIGURE 1 Histogram of guilt ratings in Study 2.

TABLE 2 Regression model for perceived guilt of the defendant in Study 2.

		Adj R^2	$F(3, 909)$	P	
		.028	9.64	<.001	
	Estimate	SE	t	P	Bootstrapped CI[a]
(Intercept)	52.87	0.90	58.94	<.001	(51.27, 54.54)
Affirmation condition	−1.62	1.28	−1.27	.205	(−4.10, 0.80)
Belief in God	−0.02	0.02	−1.08	.279	(−0.06, 0.02)
Affirmation condition × Belief in God	0.11	0.03	4.13	<.001	(0.06, 0.17)

Note: Predictor variables include a dummy variable denoting the affirmation condition; the observer's degree of belief in God; and the interaction of these two variables.

[a]$r = 10,000$ bootstrapped regressions, bias-corrected and accelerated (BCa) 95% confidence intervals (CI).

Religious affiliates ($n = 348$) discriminated between defendants who swore an oath and those who made an affirmation ($p < .001$), while non-affiliates ($n = 467$) did not ($p = .780$). We also examined simple effects of having a religious affiliation (vs. no affiliation) in each declaration condition. Guilt ratings in the oath condition did not differ between affiliates and non-affiliates ($p = .108$), but affiliates perceived defendants who made an affirmation as guiltier than non-affiliates did ($p = .006$).

Thus, unaffiliated individuals did not discriminate between affirmation-choosing and oath-choosing witnesses when estimating guilt (though the results from Study 1 indicate that unaffiliated individuals are inclined to view affirmation-choosing witnesses as relatively irreligious). Religious affiliates, however, discriminated against defendants who made an affirmation.

Discrete verdict analyses

Given that real-world verdicts are discrete (i.e. "guilty" or "not guilty") rather than continuous, we ran additional analyses where continuous guilt ratings were transformed into binary (guilty/not guilty) values.
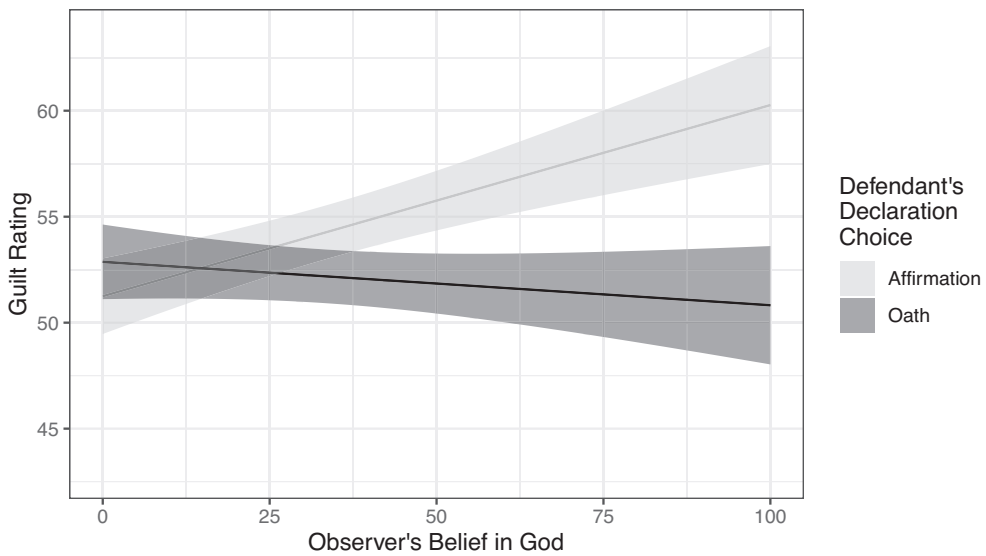
FIGURE 2   Study 2: Observer's rating of defendant's probable guilt as a function of defendant's declaration choice and observer's belief in God. *Note*: The range of the vertical axis is set as per the recommendations of Witt (2019).

TABLE 3   Regression model for perceived guilt of the defendant in Study 2.

| | | Adj R^2 | F(3, 811) | p | |
| | | .019 | 6.39 | <.001 | |
	Estimate	SE	t	p	Bootstrapped CI[a]
(Intercept)	52.55	0.92	57.02	<.001	(50.92, 54.35)
Affirmation condition	0.36	1.29	0.28	.780	(−2.04, 2.74)
Religious affiliation	−2.27	1.41	−1.61	.108	(−5.08, 0.51)
Affirmation condition × Religious affiliation	6.05	1.98	3.06	.002	(2.16, 10.04)

Note: Predictor variables include dummy variables denoting the affirmation condition and the observer having a religious affiliation, and the interaction of these two dummy variables.

[a]r = 10,000 bootstrapped regressions, bias-corrected and accelerated (BCa) 95% confidence intervals (CI).

Participants who answered 0–49 on the continuous probable guilt scale were assigned 0 (not guilty) and those who answered 51–100 were assigned 1 (guilty); "fence-sitters" (those responding 50) were randomly assigned to 0 or 1.

We then ran chi-square and binary logistic regression analyses to reexamine H1 and H2 using this binary dependent variable. Given the huge proportion of fence-sitters (57% of the sample), these analyses vary substantially depending on how the random assignment of these individuals to a verdict plays out. Accordingly, we ran each analysis 10,000 times, with fence-sitters assigned anew in each iteration.

The odds of a guilty verdict for a defendant who chose the affirmation were greater than for a defendant who chose the oath (odds ratio > 1) in 92% of cases. However, this association between verdict and declaration condition was statistically significant (χ^2 p < .05) in only 13% of the analyses.[8] As for the logistic regression analyses, including declaration condition, perceiver belief in God and their interaction produced a significant improvement in the fit of the model in 70% of the analyses. The odds ratio for the interaction term was greater than 1 in 99.9% of cases.

---

[8]In all of these cases the odds of a guilty verdict were greater for a defendant who chose the affirmation than for a defendant who chose the oath (i.e., there were no analyses where the odds of a *not* guilty verdict were significantly greater for a defendant who chose the affirmation than for a defendant who chose the oath).
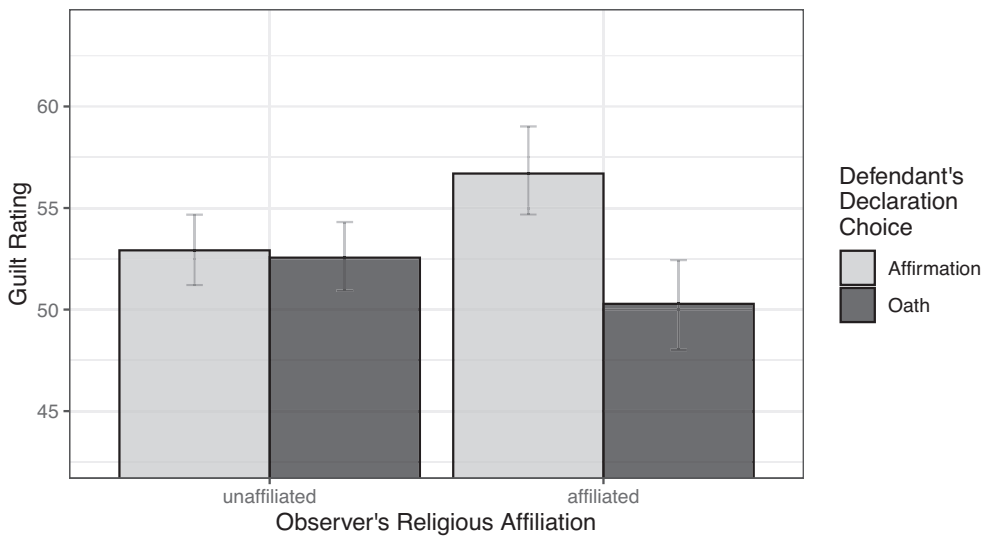
FIGURE 3 Study 2: Observer's rating of defendant's probable guilt as a function of defendant's declaration choice and observer's religious affiliation. *Note*: 1. Error bars are BCa 95% confidence intervals (CI) based on 10,000 resamples. 2. The range of the vertical axis is set as per the recommendations of Witt (2019).

The threshold for a guilty verdict in these pre-registered discrete analyses might seem somewhat low, given that a familiar standard of proof for a guilty verdict is proof "beyond a reasonable doubt" (the so-called "BARD" standard). Judges and legal scholars tend to place the BARD threshold at or near 90% certainty of the defendant's guilt (Walen, 2015). Accordingly, we ran an exploratory analysis using a 90% cut-off to generate discretised verdicts: participants who gave a rating of 90% or above were assigned 1, and those who gave a rating of less than 90% (including all of the fence-sitters) were assigned 0. The association between verdict and declaration condition was statistically significant using this cut-off, $\chi^2(1) = 4.3$, $p = .038$. The odds of a guilty verdict for a defendant who chose the affirmation were 2.49 (95% CI = [0.97, 7.17]) times higher than for a defendant who chose the oath (there was no evidence, however, that this effect was moderated by perceiver belief in God).

INTERIM DISCUSSION

The results of our first two studies indicate that court witnesses who swear an oath are, on average, much more religious than those who choose to affirm; that witnesses who swear are *perceived* as much more religious than those who affirm; that people associate choice of the oath with credible testimony; and crucially, that participants, especially religious believers and affiliates, *discriminate against hypothetical defendants who take the secular affirmation*. The latter effect is small, and does not imply that taking the affirmation instead of the oath could have a major impact on trial outcomes. Nonetheless, although effects were slightly less robust when we analysed a dichotomous guilt variable (as has been found with racial bias; see Mitchell et al., 2005), the biases we report could potentially tip the balance in cases that could go either way. Moreover, the stakes here are high, so this seems to us a paradigm example of a situation where small effects may have substantial practical implications (Cortina & Landis, 2009; Funder & Ozer, 2019; Kang et al., 2012).

An important limitation of our results to this point, however, is that the vignette we used in Study 2 was very sparse: aside from his age, nationality and gender, the only information we provided to participants about the defendant was the declaration option he had chosen. This may have made the aims of our research highly transparent, in which case demand characteristics may have played a role. On

the other hand, we employed a between-subjects design, which is inherently more conservative and less susceptible to demand effects than a within-subjects design (Charness et al., 2012). In any case, recent empirical work indicates that online survey experiments are robust to experimenter demand. Mummolo and Peterson (2019) replicated a range of experimental designs and showed that providing participants with information about experimenter expectations did not alter the treatment effects in these experiments – even financial incentives to respond in accordance with these expectations failed to consistently induce demand effects.

Nevertheless, additional data from a more elaborate and ecologically valid paradigm is needed to provide a better estimate of the real-world effect of declaration choice. Accordingly, in our third study we embedded our experimental manipulation in a much more detailed courtroom scenario. We reasoned that doing so could shed light on an interesting discrepancy between our Study 2 results and previous work on intuitive atheist distrust: whereas earlier work had suggested that even non-believers harbour moral prejudice against atheists (Gervais et al., 2017), we found no evidence that religious non-affiliates discriminated against defendants who chose a secular affirmation. It could be that most people (religious or atheist) harbour intuitive distrust about atheists, but that atheists are more inclined to consciously override these intuitions when possible. We reasoned that embedding our experimental manipulation of declaration choice in a detailed trial scenario, containing a range of excerpts from the courtroom protocol, should render it less salient, in which case activated intuitions might be more difficult to consciously override. If so, both religious *and* non-religious participants might exhibit declaration-based discrimination (though we still expected such discrimination to be stronger among religious believers).

STUDY THREE

The approved Stage 1 protocol for Study 3 (comprising pre-registered hypotheses, design, data collection, and analysis plan) can be accessed on the Open Science Framework (OSF) at https://osf.io/a35sv. A de-identified data file and analysis script for the study is also available on the OSF at https://osf.io/rk8ds/?view_only=None.

Research questions

The purpose of our third study was to investigate whether the type of legal declaration (religious vs. secular) made by defendants in a trial can influence trial outcomes. This final experiment forms the primary basis for our conclusions on this matter.

We registered three hypotheses:

H1: Mock jurors would be more likely to find defendants guilty if the defendants had chosen to make an affirmation than if they had chosen to swear an oath.
H2: The effect in H1 would be moderated by mock-juror belief in God, with declaration-based discrimination being stronger among believers.
H3: Mock jurors who chose the oath themselves (when being sworn in) would believe in God more strongly than those who chose to affirm.

Method

Overview of design

Participants acted as jurors in an animated mock trial. We manipulated the trial information to vary the declaration made by the defendant and randomly allocated participants to one of two defendant

declaration choice conditions: *oath* or *affirmation*. Mock jurors were asked to render a verdict for the defendant, and to indicate their confidence in this verdict.

Participants

Power analysis and planned sample size

We required 1524 participants to attain 90% power ($a = .05$) to detect a small effect ($d = 0.15$, as per Study 2) in one of our primary analyses (one-tailed independent samples t-test[9]). We planned to over-recruit by about a third to account for pre-registered participant exclusions (see Data exclusions section below), leaving a final planned $N = 2040$.

Recruitment strategy

We again recruited participants via *Prolific*. We screened out any participants who took part in Studies 1 or 2. To avoid duplicate submissions, we enabled the "Prevent Multiple Submissions" option in *Qualtrics*. To ensure a good spread of religious belief among our participants, we simultaneously launched four identical studies on *Prolific* ($N = 510$ per study), using *Prolific's* custom prescreening facility to target the following four mutually exclusive groups[10]:

1. Participants who selected *Non Religious* when asked "What is your religious affiliation?" AND who selected *Atheist* when asked "Which of the following do you most identify as?"
2. Participants who selected *Non Religious* when asked "What is your religious affiliation?" AND who selected *Agnostic* when asked "Which of the following do you most identify as?"
3. Participants who selected any of *Buddhism, Christianity, Hinduism, Islam, Judaism* or *Sikhism* when asked "What is your religious affiliation?" AND who selected *None/Rather not say* when asked "Do you participate in regular religious activities?"
4. Participants who selected any of *Buddhism, Christianity, Hinduism, Islam, Judaism* or *Sikhism* when asked "What is your religious affiliation?" AND who selected either *Yes. Both public and private, Yes. Public only* or *Yes. Private only* when asked "Do you participate in regular religious activities?"

Data collection

Data collection commenced (for all four studies simultaneously) at 10 am BST on Monday 11th July 2022, and was completed by 7:30 pm the same day. Participants were paid £2.00 for a survey advertised as taking about 15 min to complete. The median completion time was 14.5 min, and participants were paid £8.07 per hour on average, which *Prolific* designates a "Good" rate of pay. Though all studies successfully recruited $N = 510$ participants, the data of nine participants did not appear in the raw Qualtrics data file, leaving an initial $N = 2031$. Two participants noted that they had entered incorrect responses by accident, one mistakenly entering their age and one accidentally clicking "I don't know" when asked about the declaration chosen by the defendant. We replaced these responses with their intended responses. Ethical approval was obtained through the self-certification process at Royal Holloway, University of London, and all participants provided informed consent at the outset. Participants were informed they could withdraw their data after completion of the study, provided they contacted us within a week (none did so).

[9]To test H1, we analysed the effect of defendant declaration choice firstly on binary verdicts using Pearson's chi-square test, and secondly on a continuous measure of guilt (derived from verdicts combined with confidence), using a t-test. We based our power analysis on the latter as a comparable effect size was available from our Study 2.

[10]A number of studies of *Mechanical Turk* (*MTurk*) samples have shown that such samples are skewed toward nonreligion, with a disproportionately high number of *MTurk* workers identifying as atheists (Burnham et al., 2018; Casey et al., 2017; Levay et al., 2016; Lewis et al., 2015). Assuming this bias is not unique to *MTurk*, but is a more general characteristic of crowdsourcing marketplace platforms, it seems reasonable to apply *Prolific* prescreening criteria to target participants across a range of religiosity.

Data exclusions

We excluded 210 participants (10.3% of the sample), as follows:

- 2 participants who skipped a response (one omitted to supply their age and the other omitted to indicate their belief in God).
- 197 participants who failed either or both of the comprehension questions about witness responses (see below). A further 210 participants failed an additional question about the declaration chosen by the defendant (again, see below), but for our primary analyses we retained these participants.[11]
- In addition to our pre-registered exclusions, we excluded a further 11 participants: 5 participants who experienced technical glitches which meant they watched the trial video more than once, and 6 participants who reported problems hearing the video or reading the text. Excluding these participants did not change any of our results.
- No participants took the survey more than once (as indicated by duplicate *Prolific* IDs).

Characterizing the sample

The final sample comprised 1821 participants, well over the 1524 stipulated by our power analysis. The breakdown of self-reported gender was as follows: 911 females (50% of the sample[12]), 901 males, 6 non-binary individuals, 2 who preferred to self-describe and 1 who preferred not to disclose their gender. The age range was 18–75[13] (mean = 41.2, *SD* = 13.4). The breakdown of self-reported religious affiliation was as follows: 712 Christians, 51 Muslims, 19 Hindus, 13 Buddhists, 13 Jews and 9 Sikhs. Three hundred and eighty-eight participants identified as Atheist and 308 indicated they had no affiliation. All participants were British citizens and residents with fluency in English.

Materials and procedure

Participants watched an "animated" audiovisual version of a trial transcript and acted as mock jurors. The transcript we employed has been used in multiple previous studies with North American samples (Hunt & Budesheim, 2004; Maeder & Hunt, 2011; Maeder & McManus, 2020). The case involves a defendant who is charged with robbery. The transcript includes opening and closing statements from the barristers, as well as direct testimony from – and cross-examination of – the arresting officer, a witness, one of the alleged victims of the robbery, and the defendant. Maeder and McManus (2020) showed that in the absence of manipulations, this transcript produces an even split of verdicts. We modified the original transcript so that it fit a British setting, changing names of people and places as well as dates. The crime in the original transcript was assault and robbery, but we removed any reference to the assault to avoid participants having to render a verdict for the defendant on both charges.

In order to embed our oath/affirmation manipulation in this transcript, we made appropriate further modifications. To do this, we drew inspiration from a different transcript, produced by the Courts Administration Authority of South Australia (R vs. Hopper & Lush), which embeds a series of comments and explanations of courtroom protocol in the narrative (happily, also including comments about the oath/affirmation distinction, which we adapted for our purposes). Participants were randomly assigned to one of two versions of our animated transcript (with the constraint that groups were equal in size), one of which described the defendant as taking an oath, the other of which described the defendant as making an affirmation (for the full written transcript see the Supporting Information).

Participants were "invited to participate in a study about how we process information in a legal context." Rather than having participants simply read the transcript, we presented an "animated"

---

[11]We subsequently repeated all planned analyses after excluding any participants who failed this question. Results were in general virtually identical when excluding these participants, although we note one instance below where a key difference approached significance ($p \leq .05$) with these participants excluded.

[12]We used *Prolific's* "Balance sample" feature to ensure an equal distribution of male and female participants in each survey.

[13]To qualify for jury service in Britain, a person must be at least 18, and under 76, years of age.

version of the transcript, produced by combining photographs of each witness with an audio recording of their dialogue (with different voices reading the different players, and accompanying subtitles). Participants who consented were asked to watch the animated trial, imagining that they were a member of the jury.

Before the animation began, we "swore in" our participants as mock jurors:

> You will now watch an edited video of a trial. You will be asked to play the role of a juror. To make this as realistic as possible, we ask that you swear or affirm (just as you would in a real trial) that you will consider the issues faithfully, according to the evidence. To proceed, please select one of the following declarations (in a real trial they have the same legal effect):

- I swear by Almighty God that I will faithfully try the defendant and give a true verdict according to the evidence.
- I solemnly, sincerely and truly declare and affirm that I will faithfully try the defendant and give a true verdict according to the evidence.[14]

We think that asking our participants to themselves choose a declaration had several virtues. First, this feature is ecologically valid, as jurors in a real trial are themselves sworn in before the trial commences. Second, doing this allowed us to present the oath/affirmation distinction in a way that was more engaging, and less suspicious, than extended editorializing at the point in the trial where the defendant was sworn in. Third, having our participants promise they would consider the issues faithfully may have boosted their attention and comprehension. Finally, we reasoned that collecting data about the declaration participants themselves chose might give us extra traction on our research questions. In particular, it could allow us to distinguish declaration-specific effects (wherein people judge defendants who swear and affirm differently because of what they infer about their religiosity and credibility) from effects of identity and group membership (wherein people interpret declaration choice as a sign of group membership, and judge ingroup members and outgroup members differently).

Once the animation concluded, participants were asked to render a verdict for the defendant using a dichotomous guilty/not guilty measure,[15] before rating their confidence in this verdict. Confidence ratings were made on a slider, anchored by labels reading "not at all confident" and "very confident". We used a scale from 0 to 10, with these values hidden from participants.

To check their attention and comprehension, participants then answered three multiple-choice questions about the transcript content (two questions about witness responses and one question about the declaration chosen by the defendant), before providing their age, gender, religious affiliation and belief in God (the latter on a 0–100 scale). Finally, for exploratory purposes, we administered the six-item Very Short Authoritarianism scale (Bizumic & Duckitt, 2018).[16] Throughout the survey, we requested a response to any question participants attempted to skip but allowed them to proceed without answering.

---

[14]We counterbalanced whether the oath or affirmation option was presented first.

[15]An initial pilot study in which participants read the transcript ($N = 50$) revealed similar results to Maeder and McManus's (2020) pilot: 23 of 50 participants gave a guilty verdict, and 27 gave a not guilty verdict. Animating the transcript, however, seemed to make the defendant more sympathetic: in a subsequent pilot study with our animated version ($N = 20$), only 3 out of 20 participants returned a guilty verdict. To remedy this, we made some minor tweaks to the footage, editing out the defence re-direct and removing some other minor lines favourable to the defendant. A final pilot ($N = 20$) indicated that these edits had had the intended effect, as verdicts were now perfectly split (10 Guilty, 10 Not Guilty). The transcript in the Supporting Information is a faithful representation of the videos we presented to participants in the main study.

[16]In developing this scale, Bizumic and Duckitt (2018) asked their participants to indicate agreement/disagreement with each of the items on a nine-point scale, ranging from *very strongly disagree* to *very strongly agree*. To optimize the measure for participants taking the survey on a mobile device, we instead used a five-point response format, ranging from *strongly disagree* to *strongly agree*.

Results

Summary statistics for key measured variables

Mock-juror declaration choice
28.2% of the mock jurors chose to swear an Oath before watching the trial, and 71.8% chose to make an Affirmation.

Verdict
Mock juror verdicts were distributed fairly equally between guilty (47.4%) and not guilty (52.6%).

Perceived guilt of the defendant
To compute this variable, we multiplied verdict confidence (a continuous variable, 0–10) by −1 for participants who gave a verdict of not guilty and by +1 for participants who gave a verdict of guilty, thus creating a 21-point scale ranging from −10 (high confidence in a Not Guilty verdict) to +10 (high confidence in a Guilty verdict). Mean = 0.7, SD = 6.8. See Figure 4 for histograms of this variable split by defendant declaration condition. In both cases, the distribution is clearly bimodal.

Mock-juror belief in God
Mean belief in God for the mock jurors was 33.5 (*SD* = 35.5). The distribution of this variable reveals three pronounced spikes, at 0, 50 and 100 (see Figure 5).

Planned analyses

Overall, defendants who affirm are not considered guiltier than those who swear an oath, nor does mock-juror belief in God moderate this effect
There was no significant overall association between the defendant's chosen declaration and the verdict returned by the participants, $\chi^2(1) = 0.23$, $p = .633$ (see Table 4). Likewise, the defendant was not
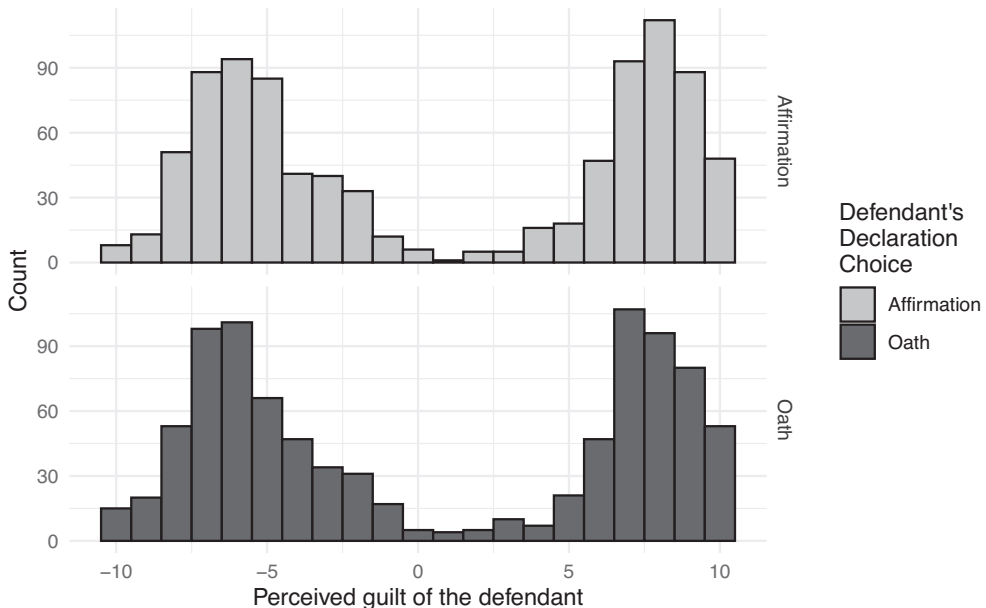

FIGURE 4   Histogram of perceived defendant guilt in Study 3.
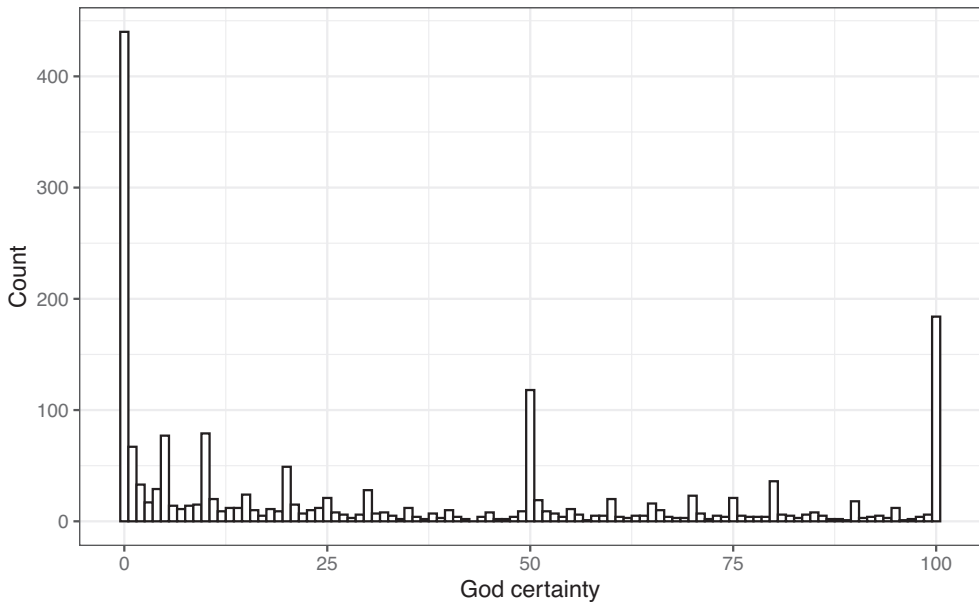
FIGURE 5   Histogram of belief in God in Study 3.

significantly perceived as guiltier overall when choosing to affirm (mean guilt rating = 0.78, $SD = 6.79$, $n = 904$) than when choosing to swear an oath (mean = 0.52, $SD = 6.86$, $n = 917$), $t(1818.97) = 0.81$, $p = .209$, $d = 0.04$; bootstrapped t-test with 10,000 resamples, BCa 95% CI for mean difference: [−0.28, ∞]; BCa 95% CI for d estimate: [−0.06, 0.13], although this result approached significance when participants who failed the question about the defendant's declaration were excluded: (mean guilt rating in Affirmation condition = 0.83, $SD = 6.79$, $n = 815$; mean rating in Oath condition = 0.33, $SD = 6.83$, $n = 796$); $t(1607.65) = 1.46$, $p = .072$, $d = 0.07$; bootstrapped t-test with 10,000 resamples, BCa 95% CI for mean difference: [−0.06, ∞]; BCa 95% CI for d estimate: [−0.03, 0.17].

There was no significant interaction between the defendant's chosen declaration and mock-juror belief in God with respect either to the juror's verdict, or to juror perceptions of the defendant's guilt (see Table 5).

Mock jurors who themselves swear an oath discriminate against defendants who affirm
We conducted a comprehensive model exploration, analysing the 32 models in Table 6 for each of the two criterion variables *verdicts* and *perceived guilt of the defendant*. These 32 models comprised every possible combination of the following five variables:

- Mock-juror declaration choice
- Defendant declaration choice

TABLE 4   Contingency table showing Study 3 participants' verdicts as a function of the defendant's chosen declaration.

	Participant's verdict		
	Guilty	**Not guilty**	**Total**
Defendant's chosen declaration			
Oath	430	487	917
Affirmation	434	470	904
Total	864	957	1821

TABLE 5 Regression model for perceived guilt of the defendant in Study 3.

| | | Adj R^2 | $F(3, 1817)$ | p | |
| | | −.000 | 0.89 | .447 | |
	Estimate	SE	t	p	Bootstrapped CI[a]
(Intercept)	0.41	0.31	1.34	.181	(−0.18, 1.00)
Defendant affirmation	0.09	0.44	0.20	.843	(−0.75, 0.93)
Juror belief in God	0.00	0.01	0.53	.594	(−0.01, 0.02)
Defendant affirmation × Juror belief in God	0.01	0.01	0.56	.577	(−0.01, 0.02)

Note: Predictor variables include a dummy variable denoting defendant choice of the affirmation; the mock juror's degree of belief in God; and the interaction of these two variables.

[a]$r = 10,000$ bootstrapped regressions, bias-corrected and accelerated (BCa) 95% confidence intervals (CI).

- Mock-juror belief in God
- The interaction between defendant declaration choice and mock-juror declaration choice
- The interaction between defendant declaration choice and mock-juror belief in God

Separately for each criterion variable, we then used the *performance* package in R (Lüdecke et al., 2021) to rank the 32 models according to their overall performance.[17] The highest performing model in each case was model 9 (to illustrate, see Table 7 for *perceived guilt of the defendant*). This model – in common with seven of the next eight highest-ranked models in each case – contains a significant interaction between *defendant declaration choice* and *mock-juror declaration choice*.[18] To probe this interaction, we regressed *verdict* on *defendant declaration choice* at each level of the *mock-juror declaration choice* variable. Mock jurors who themselves swore an oath ($n = 514$) discriminated between defendants who swore and those who affirmed (finding the latter guilty at a higher rate, $p = .018$), while mock jurors who themselves affirmed ($n = 1307$) did not discriminate ($p = .332$). Likewise, mock jurors who themselves swore an oath perceived defendants who affirmed as guiltier than defendants who swore ($p = .011$), while mock jurors who themselves affirmed did not discriminate between defendants who swore versus affirmed ($p = .453$; see Figure 6).

Mock jurors who chose the oath themselves (when being sworn in) believed in God more strongly than those who chose to affirm

Consistent with Study 1, participants who chose the oath when being sworn in reported a much stronger belief in God (Mean = 59.83, $SD = 34.52$, $n = 514$) than those who chose the affirmation (Mean belief = 23.18, $SD = 30.2$, $n = 1307$), $t(838.63) = 21.1$, $p < .001$, $d = 1.16$; bootstrapped t-test with 10,000 resamples, BCa 95% CI for mean difference: [33.79, ∞]; BCa 95% CI for d estimate: [1.04, 1.3].

Exploratory analyses

We conducted some exploratory analyses using the Very Short Authoritarianism scale (Bizumic & Duckitt, 2018). The mean score on this 6-item scale (computed by averaging responses to individual items)

---

[17]This calculation is based on normalizing all indices (including Akaike's Information Criterion, the Bayesian Information Criterion, R^2, adjusted R^2, root mean squared error, and the residual standard deviation), and taking the mean value of all indices for each model. Note that none of our models showed evidence of problematic multicollinearity: across all models we tested, the highest Variance Inflation Factor was 2.42.

[18]Note that many of our 32 models violate the so-called "principle of marginality", which holds that if interactions are included in a regression model, then all lower order main effects should be included in the model as well. However, there is some debate about the force of this principle; for example, Heathcote and Matzke (2023) discuss its limits, and state that, "to a priori rule out [an interaction-only] model because it is incomplete, results in an inferential framework incapable of discovering psychologically interesting findings" (p. 32). Our view is that each of the 32 models we tested has a clear interpretation, and we did not see any good a priori reason for excluding any of these models. Nevertheless, we ran two additional model explorations (one for *verdicts*, one for *perceived guilt of the defendant*), each on a subset of the 32 models, excluding any models that violate the principle of marginality (i.e. excluding models 5, 6, 9, 10, 12, 13, 14, 15, 16, 19, 20, 21, 22, 24, 25, 26, 29, 30 and 31). In each case, the highest performing model was model 18, which contains the (significant) interaction between *defendant declaration choice* and *mock-juror declaration choice*, as well as the two associated main effects. So, ultimately none of our conclusions are affected one way or the other by this issue.

TABLE 6 Model specifications for comprehensive model exploration in Study 3.

Model	Mock-juror declaration choice	Defendant declaration choice	Mock-juror belief in God	Defendant declaration choice × Mock-juror declaration choice	Defendant declaration choice × Mock-juror belief in God
1	(intercept only)				
2	*				
3		*			
4			*		
5				*	
6					*
7	*	*			
8	*		*		
9	*			*	
10	*				*
11		*	*		
12		*		*	
13		*			*
14			*	*	
15			*		*
16				*	*
17	*	*	*		
18	*	*		*	
19	*		*	*	
20		*	*	*	
21	*	*			*
22	*		*		*
23		*	*		*
24	*			*	*
25		*		*	*
26			*	*	*
27	*	*	*	*	
28	*	*	*		*
29	*	*		*	*
30	*		*	*	*
31		*	*	*	*
32	*	*	*	*	*

* indicates which predictor variables were included in each model.

was 2.62 ($SD = 0.7$). Cronbach's alpha was .79, indicating acceptable internal consistency. *Authoritarianism* predicted juror declarations (jurors scoring higher on authoritarianism were more likely to take the oath than the affirmation, $p < .001$, odds ratio = 3.06) and also predicted verdicts (jurors scoring higher on authoritarianism were more likely to find the defendant guilty than not guilty, $p < .001$, odds ratio = 1.47). We tried regressing *perceived guilt of the defendant* on *mock-juror declaration choice*, *defendant declaration choice*, *authoritarianism* and the two- and three-way interactions between these variables. The resulting model was significant ($F[7, 1813] = 6.28$, adjusted $R^2 = .02$, $p < .001$), as was the three-way interaction ($B = -2.22$, $SE = 1.01$, $t = -2.2$, $p = .028$, BCa 95% CI for coefficient: [-4.21, -0.15]).

To illustrate this interaction, we median-split the authoritarianism variable to create *low juror authoritarianism* and *high juror authoritarianism* subsets of the data. We then regressed *perceived guilt of the defendant*

TABLE 7 Highest performing model (model 9) for the *perceived guilt of the defendant* criterion variable in Study 3.

| | | Adj R^2 | $F(2, 1818)$ | P | |
| | | .008 | 8.06 | <.001 | Bootstrapped CI[a] |
	Estimate	*SE*	*t*	P	
(Intercept)	0.65	0.16	4.08	<.001	(0.33, 0.95)
Juror affirmation	−1.06	0.35	−2.99	.003	(−1.77, −0.33)
Juror affirmation × Defendant affirmation	−1.88	0.71	−2.65	.008	(−3.26, −0.42)

Note: Predictor variables include a centred dummy variable denoting mock-juror choice of the affirmation; and the interaction of this variable with a centred dummy denoting defendant choice of the affirmation.

[a]$r = 10,000$ bootstrapped regressions, bias-corrected and accelerated (BCa) 95% confidence intervals (CI).
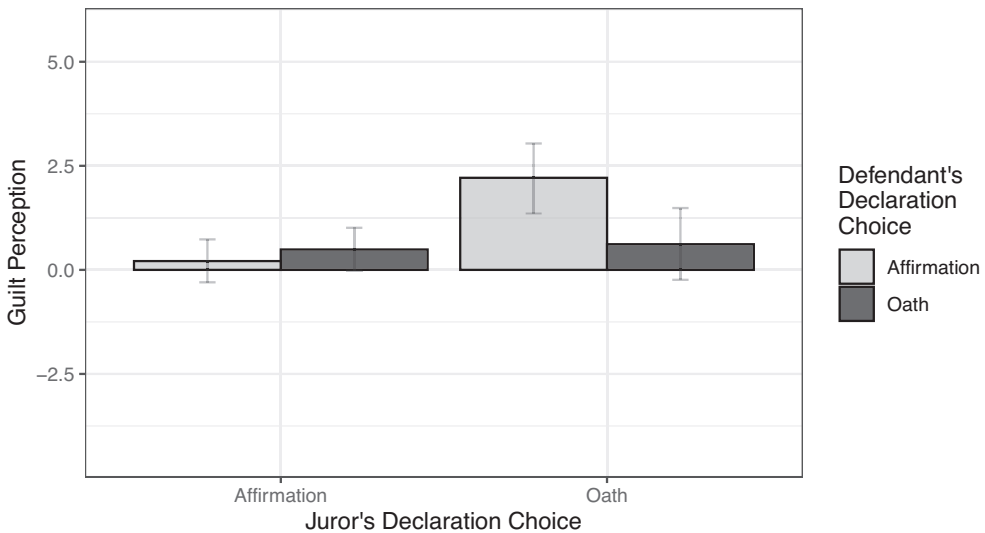
FIGURE 6 Study 3: Mock juror perception of defendant's guilt as a function of defendant declaration choice and mock juror declaration choice. *Note*: 1. Error bars are BCa 95% confidence intervals (CI) based on 10,000 resamples. 2. The range of the vertical axis is set as per the recommendations of Witt (2019).

on *mock-juror declaration choice* and *defendant declaration choice* separately for each of these subsets. For low-authoritarian jurors, there was no interaction between *defendant declaration choice* and *mock-juror declaration choice*, but for high-authoritarian jurors this interaction was highly significant ($p = .003$). High-authoritarian jurors who themselves swore an oath perceived defendants who affirmed as guiltier than defendants who swore ($p < .001$), while high-authoritarian jurors who themselves affirmed did not discriminate between defendants who swore versus affirmed ($p = .66$; see Figure 7).

GENERAL DISCUSSION

The notion that belief in God is a precondition for morality has a long history, having been articulated through the ages by philosophers, novelists and politicians (McKay & Whitehouse, 2015). George Washington, for instance, cautioned in his farewell address against "indulg[ing] the supposition that morality can be maintained without religion" (Avlon, 2017, p. 151). Moral prejudice against atheists continues to
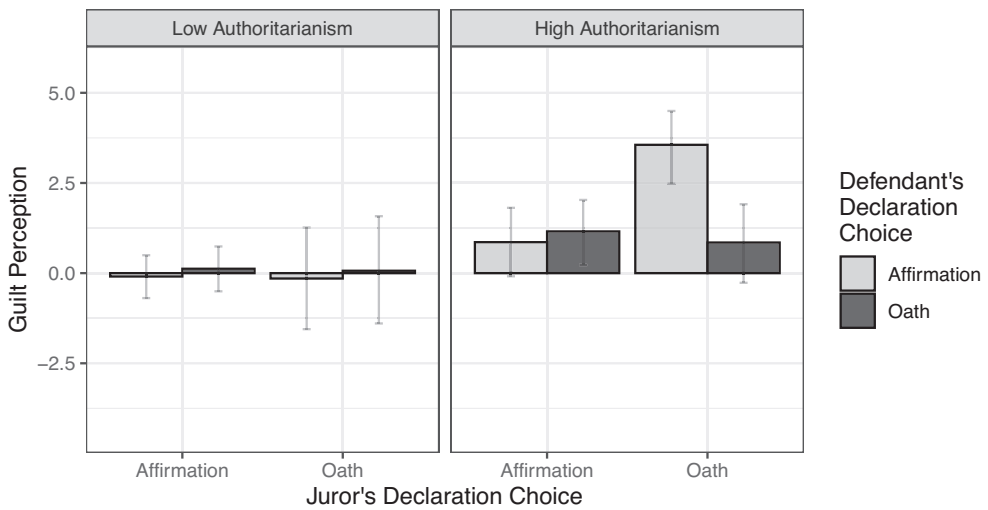
FIGURE 7 Study 3: High− and low−authoritarian jurors' perceptions of defendant's guilt as a function of defendant declaration choice and mock juror declaration choice. *Note*: 1. Error bars are BCa 95% confidence intervals (CI) based on 10,000 resamples. 2. The range of the vertical axis is set as per the recommendations of Witt (2019).

be freely expressed by those occupying,[19] or seeking to occupy,[20] the world's highest office, and 40% of American voters admit they would not vote for an otherwise well-qualified presidential candidate if he or she were an atheist (McCarthy, 2019). Though the right to decline to swear an oath is enshrined in the US Constitution, in US history only a single president, Franklin Pierce, has chosen to affirm rather than swear when being administered the presidential oath of office.[21]

The US presidency is, however, a rather rarefied domain. In the present studies, we explored the potential for discrimination against atheists in a situation many ordinary citizens will face: the courtroom. Our primary aim was to investigate whether the type of legal declaration (religious vs. secular) made by trial defendants could affect their prospects for justice. In particular, might defendants who opt to swear an oath when giving evidence enjoy more favourable judgements and outcomes than those who choose to affirm?

One objection we have faced repeatedly when suggesting that the oath/affirmation distinction could bias legal decisions is that affirming may not be a sign of religious disbelief, because some witnesses (e.g. Quakers) affirm for religious reasons. However, Study 1 indicated – and Study 3 confirmed – that witnesses who choose to swear an oath are, on average, far more religious than those who choose to affirm. Notwithstanding the historical origins of the affirmation, therefore, the affirmation appears to be a reliable signal of religious disbelief – and is perceived as such. Given prevailing negative media portrayals of religious disbelievers (van der Veen & Bleich, 2021), coupled with cross-cultural evidence of moral prejudice against them (Gervais et al., 2017; cf. Moon et al., 2021), this finding should be cause for some concern. But do jurors actually discriminate against defendants who choose to affirm?

While the results of Study 2 suggested that they do, this study employed a very sparse vignette, with the information about the hypothetical defendant's chosen declaration being highly salient. Data from a more elaborate and ecologically valid paradigm were needed to better approximate the real-world

---

[19]"I don't know that atheists should be considered citizens, nor should they be considered patriots. This is one nation under God" ~ George H. W. Bush (Krassner, 2008).

[20]"Any president who doesn't begin every day on his knees isn't fit to be commander in chief of this country" ~ Senator Ted Cruz of Texas, Republican presidential candidate (Lopez, 2015).

[21]Pierce did not choose to affirm because he was an atheist, but because he believed the recent death of his son was a punishment for his sins (Vinciguerra, 1985). Herbert Hoover, a Quaker, is often reported as having chosen to affirm rather than swear, but newsreel of his inauguration disproves this (Bendat, 2012).

influence of declaration choice. This is what we sought to provide in our main study, Study 3. We produced an engaging, audiovisual adaptation of a well-established trial transcript and recruited participants to act as mock jurors. In one between-subjects version of the trial, the defendant chose to swear an oath, in the other he chose to affirm. Our jurors were themselves required to declare their intentions to try the defendant in good faith, by swearing or affirming to supply a true verdict according to the evidence.

Overall, the defendant was not considered guiltier when choosing to affirm rather than swear (although this result approached significance when jurors who failed the question about the defendant's declaration were excluded), nor did mock-juror belief in God moderate this effect. However, jurors who themselves swore an oath *did* discriminate against the affirming defendant. How can we explain this? Although these jurors reported a much stronger belief in God than jurors who affirmed, belief in God per se does not seem to be behind this effect, because there was no significant interaction between the defendant's chosen declaration and mock-juror belief in God with respect to juror verdicts or perceptions of the defendant's guilt. Of course, we did find such an interaction in Study 2 – the higher the observer's belief in God in that study, the guiltier they perceived affirming defendants to be – but the vignette we used there was very sparse, and it may be that the effect of juror belief in God is negligible when more contextual details are available to jurors. At the same time, that jurors who themselves swore an oath discriminated against affirming defendants does not seem to be a matter of simple in-group bias (i.e. jurors favouring defendants who chose the same declaration they did, whichever declaration it was), because jurors who themselves affirmed did not favour defendants who affirmed.

Our exploratory analyses with authoritarianism shed some light here, as these indicated that the juror bias in our sample was specific to high-authoritarian jurors. Among other characteristics, the authoritarian mindset is classically conceived as prejudicial and punitive (Bizumic & Duckitt, 2018), qualities certainly borne out by our Study 3 data. In addition, an integral subdimension of the authoritarian personality is Traditionalism. It may be that many high-authoritarian jurors consider the oath the traditional (and therefore correct) declaration to choose, and view with suspicion those who instead choose to affirm. In Study 1, 15% of participants who chose the oath reported doing so because they saw it as the more "traditional" alternative (cf. 0% of participants who chose the affirmation; for full details see the Supporting Information).

Real-world implications

Whatever the underlying mechanism, that a subset of jurors discriminate against defendants who affirm is cause for concern. The effect is only small (the winning model from our model comparison explained only about 1% of the variance in guilt perceptions), but recent authors (Funder & Ozer, 2019; Götz et al., 2022; cf. Anvari et al., 2022) have highlighted how small effects can have substantial consequences, particularly when considered at scale and over long time periods. In this spirit, we briefly consider the real-world implications of our findings.

In 2011, 12,152 defendants were convicted by juries in the Crown Court in England and Wales, while 5757 were acquitted by jury verdict (Ministry of Justice, 2012). How many of the 12,152 convicted defendants might have been acquitted if there had been no jurors biased against the affirmation?

Among numerous other uncertainties in this exercise, one prominent unknown is how jury deliberation might interact with individual juror biases. On the one hand, juries may evince "wisdom" (Galton, 1907; Surowiecki, 2005), such that the collective context nullifies individual biases. On the other hand, it could be that the collective jury context actually *amplifies* individual biases (Lynch & Haney, 2009[22]; Sulik et al., 2021). For the sake of argument, we assume that deliberation has no systematic effect, and we

---

[22]Lynch and Haney (2009) found that jury deliberation *exacerbated* the tendency of white mock jurors to disproportionately sentence black defendants to death.

treat trial verdicts as rendered by single jurors rather than by juries of twelve. We additionally assume that 50% of defendants in real trials swear and 50% affirm.[23]

If none of the 12,152 convicted defendants had been tried by biased (i.e. oath-taking) jurors, we could conclude that 6076 of the defendants (half of 12,152) had chosen the oath and 6076 the affirmation. If all convicted defendants had been tried by oath-taking jurors, however, we can estimate – using our finding of juror bias in Study 3 – that 5665 of the convicted defendants had chosen the oath and 6487 the affirmation,[24] in which case bias against affirming defendants would have resulted in 822 additional convictions in the space of a single year. Given that not all jurors are biased, the true number of additional convictions will be substantially less (maybe half this figure if we assume 50% of jurors in real trials are oath-takers), but still potentially in the hundreds every year.[25]

CONCLUSION

In a 2001 review of the criminal courts of England and Wales, Lord Justice Auld recommended that the witness's oath and affirmation be replaced by a solemn promise to tell the truth (Auld, 2001), a proposal that provoked protests from church leaders (MacCallum, 2001). In 2013 a proposal to abolish the oath in English and Welsh courts was debated and rejected by the Magistrates' Association (Pigott, 2013). Perhaps ironically, opponents of the proposal – again, including religious leaders – argued that the oath should be retained because it strengthens the value of witnesses' evidence. One could argue that one declaration option being perceived as a stronger signal of credibility than the other is precisely *why* the choice between them should be removed and the oath should be abolished (McKay & Davis, 2017).[26] Otherwise, non-religious defendants who choose to affirm, rather than "tell a lie" and swear an oath in bad faith, may be taking a risk – "subjecting themselves to a disability", in the words of Justice Scott. Ultimately, continued use of the oath may make justice more difficult to obtain for those who are unwilling to swear by a God they do not believe in.

---

[23]For a justification of this assumption see the Supporting Information.

[24]If we assume that 5665 convicted defendants swore, 6487 convicted defendants affirmed, 3289.5 acquitted defendants swore and 2467.5 acquitted defendants affirmed, then the odds of a guilty verdict for an affirming defendant were 1.53 times higher than those for a defendant who swore, which is the same odds ratio for oath-swearing jurors in Study 3.

[25]An additional caveat is that the evidence in our case was (by design) closely balanced, whereas in many trials the case will be more clearly "open and shut". On the other hand, as guilty verdicts require at least a 10-2 majority, there will be many cases where a single juror could have tipped the scales (In 2011, 11% of guilty verdicts in the English and Welsh Crown Court were returned by 10-2 majority; Ministry of Justice, 2012).

[26]In Ireland in 2020, a bill was passed (the Civil Law and Criminal Law [Miscellaneous Provisions] Act 2020) which removed the requirement for witnesses to swear before God or make an affirmation when filing affidavits (Donnelly, 2020). The oath and affirmation system was described by Law Society director general Ken Murphy as placing witnesses "in a position of embarrassment and indignity" and being "contrary to the right to privacy" (Gallagher, 2020). At the time of writing, jurors and witnesses in Irish courts still have to swear or affirm.

Hodgkins for voicing the characters in the animated trial; and to Michael Sillence Davis for his help with editing the images.

OPEN RESEARCH BADGES

This article has earned Open Data and Preregistered Research Designs badges. Data and the preregistered design and analysis plan are available at https://osf.io/rk8ds/?view_only=5792786a0446493cb-b0ab78f9124df22 (Open Data); Study 1: http://aspredicted.org/blind.php?x=83mr9d. Study 2: http://aspredicted.org/blind.php?x=64px4f. Study 3: https://osf.io/a35sv (Preregistered).

DATA AVAILABILITY STATEMENT

De-identified data files and analysis scripts for all three studies in this manuscript are available on the Open Science Framework: https://osf.io/rk8ds/?view_only=None. Following advice from our Institution's Data Protection Manager, we have split the data file for Study 1 in two, separating participants' stated reasons for their own declaration choices from any identifying information (age, gender, religious affiliation) and then shuffling the rows in the file containing the open-ended responses.

ORCID

Ryan T. McKay https://orcid.org/0000-0001-7781-1539
Will Gervais https://orcid.org/0000-0001-7790-1665
Colin J. Davis https://orcid.org/0000-0001-5814-5104

REFERENCES

Anvari, F., Kievit, R., Lakens, D., Pennington, C. R., Przybylski, A. K., Tiokhin, L., Wiernik, B. M., & Orben, A. (2022). Not all effects are indispensable: Psychological science requires verifiable lines of reasoning for whether an effect matters. *Perspectives on Psychological Science*, *18*, 503–507. https://doi.org/10.1177/17456916221091565

Auld, R. E. (2001). *Review of the Criminal Courts of England and Wales*. https://www.criminal-courts-review.org.uk/ccr-11.htm

Avlon, J. (2017). *Washington's farewell: The founding father's warning to future generations*. Simon & Schuster.

Bendat, J. (2012). *Democracy's big day: The inauguration of our president, 1789–2013*. iUniverse.

Bizumic, B., & Duckitt, J. (2018). Investigating right wing authoritarianism with a very short authoritarianism scale. *Journal of Social and Political Psychology*, *6*(1), 129–150. https://doi.org/10.5964/jspp.v6i1.835

Bornstein, B., & Miller, M. (2009). *God in the courtroom: Religion's role at trial*. Oxford University Press.

Boyd, C. L., Epstein, L., & Martin, A. D. (2010). Untangling the causal effects of sex on judging. *American Journal of Political Science*, *54*(2), 389–411.

Brown-Iannuzzi, J. L., Golding, J. M., Gervais, W. M., Lynch, K. R., Wasarhaley, N. E., & Bainter, S. (2021). Will jurors believe nonbelievers? Perceptions of atheist rape victims in the courtroom. *Psychology of Religion and Spirituality*, *13*(1), 119–126. https://doi.org/10.1037/rel0000278

Burnham, M. J., Le, Y. K., & Piedmont, R. L. (2018). Who is Mturk? Personal characteristics and sample consistency of these online workers. *Mental Health, Religion and Culture*, *21*(9–10), 934–944. https://doi.org/10.1080/13674676.2018.1486394

Casey, L. S., Chandler, J., Levine, A. S., Proctor, A., & Strolovitch, D. Z. (2017). Intertemporal differences among MTurk workers: Time-based sample variations and implications for online data collection. *SAGE Open*, *7*(2), 1–15. https://doi.org/10.1177/2158244017712774

Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization*, *81*, 1–8.

Cho, K., Barnes, C. M., & Guanara, C. L. (2017). Sleepy punishers are harsh punishers: Daylight saving time and legal sentences. *Psychological Science*, *28*(2), 242–247.

Cortina, J. M., & Landis, R. S. (2009). When small effect sizes tell a big story, and when large effect sizes don't. In C. E. Lance & R. J. Vandenberg (Eds.), *Statistical and methodological myths and urban legends: Doctrine, verity and fable in the organizational and social sciences* (pp. 287–308). Routledge/Taylor & Francis Group.

Danziger, S., Levav, J., & Avnaim-Pesso, L. (2011). Extraneous factors in judicial decisions. *PNAS*, *108*(17), 6889–6892.

Donnelly, J. (2020, August 11). Religious oaths must be removed from State activities. *The Irish Times*. https://www.irishtimes.com/opinion/religious-oaths-must-be-removed-from-state-activities-1.4326762

Edgell, P., Gerteis, J., & Hartmann, D. (2006). Atheists as "other": Moral boundaries and cultural membership in American society. *American Sociological Review*, *71*(2), 211–234.

Englich, B., Mussweiler, T., & Strack, F. (2006). Playing dice with criminal sentences: The influence of irrelevant anchors on experts' judicial decision making. *Personality and Social Psychology Bulletin*, *32*(2), 188–200. https://doi.org/10.1177/0146167205282152

Eren, O., & Mocan, N. (2018). Emotional judges and unlucky juveniles. *American Economic Journal: Applied Economics*, *10*(3), 171–205.

Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science*, *2*(2), 156–168.

Gallagher, C. (2020, July 22). 'Embarrassing' system of oaths and affirmations to be abolished. *The Irish Times*. https://www.irish-times.com/news/crime-and-law/embarrassing-system-of-oaths-and-affirmations-to-be-abolished-1.4310299

Galton, F. (1907). Vox populi. *Nature*, *75*, 450–451.

Gervais, W. M. (2011). Finding the faithless: Perceived atheist prevalence reduces anti-atheist prejudice. *Personality and Social Psychology Bulletin*, *37*, 543–556.

Gervais, W. M. (2013). In godlessness we distrust: Using social psychology to solve the puzzle of anti-atheist prejudice. *Social and Personality Psychology Compass*, *7*, 366–377.

Gervais, W. M. (2014). Everything is permitted? People intuitively judge immorality as representative of atheists. *PLoS One*, *9*(4), e92302.

Gervais, W. M., & Norenzayan, A. (2012). Reminders of secular authority reduce believers' distrust of atheists. *Psychological Science*, *23*, 483–491.

Gervais, W. M., & Norenzayan, A. (2013). Religion and the origins of anti-atheist prejudice. In S. Clarke, R. Powell, & J. Savulescu (Eds.), *Religion, intolerance and conflict: A scientific and conceptual investigation* (pp. 126–145). Oxford University Press.

Gervais, W. M., Shariff, A. F., & Norenzayan, A. (2011). Do you believe in atheists? Distrust is central to anti-atheist prejudice. *Journal of Personality and Social Psychology*, *101*, 1189–1206.

Gervais, W. M., Xygalatas, D., McKay, R. T., van Elk, M., Buchtel, E. E., Aveyard, M., Schiavone, S. R., Dar-Nimrod, I., Svedholm-Häkkinen, A. M., Riekki, T., Klocová, E. K., Ramsay, J. E., & Bulbulia, J. (2017). Global evidence of extreme intuitive moral prejudice against atheists. *Nature Human Behaviour*, *1*, 0151.

Giddings, L., & Dunn, T. J. (2016). The robustness of anti-atheist prejudice as measured by way of cognitive errors. *The International Journal for the Psychology of Religion*, *26*(2), 124–135.

Glynn, A. N., & Sen, M. (2015). Identifying judicial empathy: Does having daughters cause judges to rule for women's issues? *American Journal of Political Science*, *59*(1), 37–54.

Götz, F. M., Gosling, S. D., & Rentfrow, P. J. (2022). Small effects: The indispensable foundation for a cumulative psychological science. *Perspectives on Psychological Science*, *17*(1), 205–215. https://doi.org/10.1177/1745691620984483

Grattan, P. R. (1847). *Reports of cases decided in the supreme court of appeals, and in the general court, of Virginia*. Shepherd and Colin.

Heathcote, A., & Matzke, D. (2023). The limits of marginality. *Computational Brain & Behavior*, *6*, 28–34. https://doi.org/10.1007/s42113-021-00120-3

Hughes, J., Grossmann, I., & Cohen, A. B. (2015). Tolerating the "doubting Thomas": How centrality of religious beliefs vs. practices influences prejudice against atheists. *Frontiers in Psychology*, *6*, Article 1352. https://doi.org/10.3389/fpsyg.2015.01352

Hunt, J. S. (2015). Race, ethnicity, and culture in jury decision making. *Annual Review of Law and Social Science*, *11*, 269–288. https://doi.org/10.1146/annurev-lawsocsci-120814-121723

Hunt, J. S., & Budesheim, T. L. (2004). How jurors use and misuse character evidence. *Journal of Applied Psychology*, *89*(347), 361.

Kang, J., Bennett, M., Carbado, D., & Casey, P. (2012). Implicit bias in the courtroom. *UCLA Law Review*, *59*, 1124–1186.

Krassner, P. (2008, January 8). There are no atheists in the White House. https://www.huffpost.com/entry/there-are-no-atheists-in_b_80441

Levay, K. E., Freese, J., & Druckman, J. N. (2016). The demographic and political composition of mechanical Turk samples. *SAGE Open*, *6*(1), 1–17. https://doi.org/10.1177/2158244016636433

Lewis, A. R., Djupe, P. A., Mockabee, S. T., & Wu, J. S.-Y. (2015). The (non) religion of mechanical Turk workers. *Journal for the Scientific Study of Religion*, *54*(2), 419–428.

Lopez, G. (2015, November 10). Ted Cruz: Atheists shouldn't be president. https://www.vox.com/2015/11/10/9699778/ted-cruz-atheists-president

Lüdecke, D., Ben-Shachar, M., Patil, I., Waggoner, P., & Makowski, D. (2021). Performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software*, *6*(60), 3139. https://doi.org/10.21105/joss.03139

Lynch, M., & Haney, C. (2009). Capital jury deliberation: Effects on death sentencing, comprehension, and discrimination. *Law and Human Behavior*, *33*, 481–496. https://doi.org/10.1007/s10979-008-9168-2

MacCallum, V. (2001, October 26). Divine intervention as Auld vows to abolish oath. *The Law Society Gazette*. https://www.lawgazette.co.uk/news/divine-intervention-as-auld-vows-to-abolish-oath/35355.article

Maeder, E. M., & Hunt, J. S. (2011). Talking about a black man: The influence of defendant and character witness race on jurors' use of character evidence. *Behavioral Sciences & the Law*, *29*(4), 608–620.

Maeder, E. M., & McManus, L. A. (2020). Mosaic or melting pot? Race and juror decision making in Canada and the United States. *Journal of Interpersonal Violence*, *37*, NP991–NP1012. https://doi.org/10.1177/0886260520917512

Maitland, F. W. (1908/2008). *The constitutional history of England: A course of lectures delivered by F. W. Maitland*. The Lawbook Exchange.

McCarthy, J. (2019, May 9). Less than half in U.S. would vote for a socialist for president. https://news.gallup.com/poll/254120/less-half-vote-socialist-president.aspx

McKay, R., & Davis, C. (2017, August 10). Abolish the oath: Moral prejudice against atheists may bias courtroom decisions. *The Conversation*. https://theconversation.com/abolish-the-oath-moral-prejudice-against-atheists-may-bias-courtroom-decisions-82230

McKay, R., & Whitehouse, H. (2015). Religion and morality. *Psychological Bulletin*, *141*(2), 447–473.

Miller, M. K., & Bornstein, B. H. (2006). The use of religion in death penalty sentencing trials. *Law and Human Behavior*, *30*, 675–684.

Miller, M. K., Maskaly, J., Peoples, C. D., & Sigillo, A. E. (2014). The relationship between mock jurors' religious characteristics and their verdicts and sentencing decisions. *Psychology of Religion and Spirituality*, *6*(3), 188–197.

Ministry of Justice. (2012). *Judicial and court statistics 2011*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/217494/judicial-court-stats-2011.pdf

Mitchell, T. L., Haw, R. M., Pfeifer, J. E., & Meissner, C. A. (2005). Racial bias in mock juror decision-making: A meta-analytic review of defendant treatment. *Law and Human Behavior*, *29*(6), 621–637.

Moon, J. W., Krems, J. A., & Cohen, A. B. (2021). Is there anything good about atheists? Exploring positive and negative stereotypes of the religious and nonreligious. *Social Psychological and Personality Science*, *12*(8), 1505–1516. https://doi.org/10.1177/1948550620982703

Mummolo, J., & Peterson, E. (2019). Demand effects in survey experiments: An empirical assessment. *American Political Science Review*, *113*(2), 517–529.

Pew Research Center. (2020). The Global God Divide. https://www.pewresearch.org/global/2020/07/20/the-global-god-divide/

Pigott, R. (2013, October 19). Motion to end Bible oaths in court defeated. https://www.bbc.co.uk/news/uk-24588854

R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. Vienna, Austria. https://www.R-project.org/

Simon, D. (2012). *In doubt: The psychology of the criminal justice process*. Harvard University Press.

Sommers, S. R., & Marotta, S. A. (2014). Racial disparities in legal outcomes: On policing, charging decisions, and criminal trial proceedings. *Policy Insights From the Behavioral and Brain Sciences*, *1*(1), 103–111.

Spamann, H. (2018). Are sleepy punishers really harsh punishers? Comment on Cho, Barnes, and Guanara (2017). *Psychological Science*, *29*(6), 1006–1009.

Sulik, J., Efferson, C., & McKay, R. (2021). Collectively jumping to conclusions: Social information amplifies the tendency to gather insufficient data. *Journal of Experimental Psychology: General*, *150*(11), 2309–2320. https://doi.org/10.1037/xge0001044

Surowiecki, J. (2005). *The wisdom of crowds*. Random House Inc.

Tan, J. H. W., & Vogel, C. (2008). Religion and trust: An experimental study. *Journal of Economic Psychology*, *29*(6), 832–848.

van der Veen, A. M., & Bleich, E. (2021). Atheism in US and UK newspapers: Negativity about non-belief and non-believers. *Religion*, *12*, 291. https://doi.org/10.3390/rel12050291

Vinciguerra, T. (1985, Jan 31). Why Pierce affirmed [Letter to the editor]. *The New York Times*, A22. https://www.nytimes.com/1985/01/31/opinion/l-why-pierce-affirmed-252205.html

Walen, A. (2015). Proof beyond a reasonable doubt: A balanced retributive account. *Louisiana Law Review*, *76*(2), 354–446.

Witt, J. K. (2019). Graph construction: An empirical investigation on setting the range of the y-axis. *Meta-Psychology*, *3*, MP.2018.895. https://doi.org/10.15626/MP.2018.895

Yamamoto, S., Maeder, E. M., Mossiere, A., & Brown, D. (2019). The influence of defendant body size and defendant gender on mock juror decision-making. *Cogent Psychology*, *6*, 1674091. https://doi.org/10.1080/23311908.2019.1674091

Zuckerman, P. (2008). *Society without god*. NYU Press.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.