

How do philosophers evaluate natural theological arguments? An experimental philosophical investigation

Helen De Cruz and Johan De Smedt

Forthcoming in Helen De Cruz & Ryan Nichols (Eds), *Advances in Religion, Cognitive Science, and Experimental Philosophy*. Forthcoming with Continuum

Note: this is a draft and deviations from the final published manuscript may still occur.

Abstract

This paper provides an empirical study of how religious belief influences the views of philosophers about natural theological arguments. Philosophers rated eight arguments for and eight arguments against theism. We find a correlation between religious belief and the perceived strength of arguments: atheists tend to find arguments against theism stronger and arguments for theism weaker; theists evaluate arguments for theism as stronger than arguments against theism. The assessments of agnostics fall between those of theists and atheists. Other factors that correlate with the perceived strength of the arguments are philosophical specialization and gender. We examine the potential role of confirmation bias, early-developed intuitions, philosophical expertise and gender differences to explain our findings.

Keywords: Natural theology, arguments for theism, arguments against theism, experimental philosophy of religion, cognitive science of religion

1. Introduction: Natural theology in philosophy of religion

Arguments for or against theism have an enduring appeal, and discussions about their merits continue to dominate debates in philosophy of religion. During the Middle Ages, God was a central topic of philosophical investigation, with authors like Anselm, Avicenna (Ibn Sīnā) and Aquinas studying God's existence and attributes through reason and arguments. Although religion no longer occupies the central place in philosophy it once held, it remains of lasting interest, with popular books on the topic of God's existence regularly making bestseller lists (e.g., Dawkins 2006; Keller 2008). Moreover, cosmological and design arguments appear in different cultures and times, for instance, in Greek and Roman antiquity, 8th-10th century Hinduism, medieval Europe and the Islamic world (see De Cruz 2014a for a review).

A large survey among professional philosophers (Bourget and Chalmers 2014)¹ reveals that theists are highly represented among philosophers of religion: 72.3 percent of philosophers of religion lean toward or accept theism, compared to 11.7 percent of philosophers outside of this field. This was the highest observed correlation between a philosophical specialization and a philosophical view ($r = .351$). Several nontheist authors (e.g., Levine 2000; Trakakis 2008; Draper and Nichols 2013) maintain that the overrepresentation of Christian theists in analytic philosophy of religion is unhealthy for the field, since they would be too much influenced by prior

beliefs when evaluating religious arguments. However, a large percentage of theists does not, by itself, reveal that arguments in philosophy of religion are biased, nor does it specify what the influence of prior religious beliefs might be.

This study looks in more detail at the relationship between theism and the appraisal of natural theological arguments, i.e., arguments that aim to establish theism or another metaphysical position through observation and reason. We aim to answer two interrelated questions: how philosophers appraise arguments for or against theism, and what factors influence their evaluation of such arguments. We will look in particular at the role of religious belief, philosophical specialization and gender.

2. Methodology

The first author conducted an Internet survey. Participants rated eight arguments for theism and eight arguments against theism (Table 1) which were presented in a randomized order for each participant. Several of these arguments have different versions (e.g., the Leibnizian and the kalām cosmological arguments). We presented the arguments in a simple, generic form, with labels such as ‘ontological argument’, ‘cosmological argument’ to make the survey accessible to philosophers with diverse specializations. Participants were asked to rate how strong they thought each argument was on a scale from 1 to 5, with 1 = very weak and 5 = very strong. The prompt was ‘Please rate the following arguments according to how strong you find them.’ Arguments were presented in randomized order for each participant. Next to the ratings, participants could indicate if they were insufficiently familiar with the argument to rate it, or decline to answer. Information about professional position, philosophical areas of specialization, age, gender and religious belief was also collected. The options for religious belief were ‘theist’, ‘atheist’ or ‘agnostic/undecided’ (henceforth referred to as ‘agnostic’).

3. Results

3.1 Descriptive statistics of the participants

Participants (N = 802) were recruited through philosophy mailing lists (e.g., Philos-L) and three philosophy blogs (Prosblogion, NewApps, Feminist Philosophers). These mailing lists and blogs are mainly read by professional philosophers and graduate students in philosophy who are the target population for this survey. The majority of respondents (82 percent) were professional philosophers, working as faculty members (32.9 percent), non-tenure track PhD holders, such as postdocs, visiting assistant professors and adjuncts (15.8 percent) or graduate students (33.3 percent). The remaining respondents were undergraduates (8 percent) or not employed in academia (10 percent). Since these responses were not statistically different from those of the target population, we included them in the analysis. 24.2 percent of the respondents were female. Although the percentage of men in this sample is high, it is in line with the gender distribution among professional philosophers: the percentage of female philosophers in tenure-track or tenured positions in the US is 19.8 percent, in the UK 22.3 percent and in Canada 25.4 percent (see Buckwalter and Stich 2014 for an overview). The participants were on average 36.5 years old (SD = 11.8 years).

In this sample, 40.5 percent self-identified as theists, 40.4 percent as atheists, and 19.1 percent as agnostics. This distribution allows for robust comparisons between what people with different religious outlooks think about natural theological arguments. The percentage of philosophers of religion in this sample is also quite high.

The most commonly reported areas of specialization² were philosophy of religion (33.8 percent), metaphysics (27.8 percent), ethics (26.8 percent), epistemology (25.8 percent), history of philosophy (22.2 percent) and philosophy of mind (19.2 percent).

3.2 Correlation between religious belief and specialization in philosophy of religion

We examined whether religious belief correlates positively with specializing in philosophy of religion. Among our respondents who listed philosophy of religion as an area of specialization, 73 percent were theists, 17 percent atheists and 10 percent agnostics. Of those who did not list philosophy of religion as an area of specialization 23.9 percent were theists, 52.4 percent atheists and 23.7 percent agnostics. Philosophers of religion and non-philosophers of religion significantly differed in their religious self-identification, $\chi^2(2, N = 802) = 180.359, p = .0001$. There is a strong correlation between religious belief and philosophy of religion as an area of specialization ($r = .401$).

While our findings replicate Bourget and Chalmers's (2014) percentage of theists in philosophy of religion (72.3 percent in their study, 73 percent here), we ended up with a higher percentage of theists among non-philosophers of religion (11.7 percent in their study, 23.9 percent here). This may be the result of self-selection: given the subject matter, theists may have had more interest in completing this survey than nontheists. Alternatively, the higher percentage of theists in this study might reflect a more diverse demographic of faculty members and students. Sociological studies have found that atheism is more prevalent in elite universities than among regular faculty members (see e.g., Gross and Simmons 2009). Bourget and Chalmers (2014) selected faculty members in leading PhD granting departments as their target participants, thereby excluding departments that are less prestigious, e.g., that only provide undergraduate or master degrees. Thus, they may have underestimated the percentage of theists in the population of professional philosophers due to their focus on elite institutions.

The percentage of theists among faculty members (tenured, tenure-track and non-tenure track) was 36.6, whereas the percentage of theists in graduate and undergraduate students amounted to 45.9; the percentage of theists among non-academic respondents was 37.5. Although this difference did not reach statistical significance, it is in line with data from Bourget and Chalmers (2014) which also indicate a higher percentage of theists among graduate students (20.9 percent), compared to faculty members (16.4 percent).

3.3 The perceived strength of religious arguments

Table 7.1 provides a summary of the arguments that were presented to the participants, with the mean score for each argument. Among the arguments for theism, those that were perceived as strongest were the cosmological argument (3.0), the argument from design (2.7) and the argument from religious experience (2.7). Among the arguments against theism, respondents judged the argument from evil to be the strongest (3.5), followed by the argument from divine hiddenness (2.8) and the argument from lack of evidence (2.8).

Arguments for theism	mean score
- Cosmological argument	3.0
- Argument from design	2.7
- Argument from religious experience	2.7
- Moral argument	2.5
- Ontological argument	2.5
- Argument from beauty	2.2
- Pragmatic argument for theism	2.2
- Argument from miracles	2.1

Arguments against theism	mean score
- Argument from evil	3.5
- Argument from divine hiddenness	2.8
- Argument from lack of evidence	2.8
- Argument from parsimony	2.5
- Argument from poor design	2.4
- Argument from inconsistent revelations	2.4
- Argument from incoherence	2.4
- Pragmatic argument for atheism	2.1

Table 7.1. Arguments presented to participants (order randomized for each participant) and their mean scores. The arguments are shown in descending order of perceived strength.

What explains the differences in ratings? One factor might be familiarity: an argument that receives more attention in the philosophical literature might be perceived as stronger. To examine this possibility, we took the total number of papers indexed for each argument in the PhilPapers database, one of the largest databases of philosophical works, containing not just journal papers but also monographs and articles in edited volumes. Since only two of the arguments against theism listed here were indexed, we concentrated on arguments for the existence of God in the survey, all of which were indexed except for the argument from beauty. The numbers of papers per argument are summarized in Table 7.2.

Arguments for theism	number of papers	mean score
- Religious experience	226	2.7
- Cosmological argument	84	3.0
- Pragmatic argument for theism	70	2.2
- Ontological argument	69	2.5
- Design argument	40	2.7
- Moral argument	31	2.5
- Argument from miracles	21	2.1

Table 7.2. Number of papers indexed in PhilPapers (accessed October 15, 2014) for each of the arguments for theism, in descending number of papers.

As can be gleaned from this dataset, there is no significant correlation between the perceived strength of the arguments and the number of papers that discussed them ($p = .387$). Thus, familiarity does not predict the perceived strength of these arguments.

3.4 Religious belief predicts perceived strength of the arguments

We examined whether theists, atheists and agnostics differ from each other in how they evaluate the natural theological arguments presented in this survey (see also De Cruz 2014b for a preliminary analysis of some of these results; and De Cruz and De Smedt 2015 for a discussion of the results for the cosmological argument). We predict that philosophers evaluate arguments that support their prior beliefs as stronger than those that disconfirm their beliefs. Given that agnostics do not have a firm opinion on the matter, their beliefs would fall somewhere between those of theists and atheists.

A total score for arguments for theism was calculated by adding the individual scores for each of the eight arguments; the total score for arguments against theism is composed of the sum of all the individual scores for each of the eight arguments in that category³. Since each individual argument could be rated from 1 to 5, the maximum score an individual participant could reach for each type of argument was 40, i.e., someone who rated all arguments for theism 5 would have a total score of 40 for the arguments for theism. The minimum score was 8, i.e., someone who rated all arguments against theism 1 would have a total score of 8 for arguments against theism.

In line with our prediction, theists rated arguments that support theism significantly higher than atheists, whereas atheists rated arguments against theism significantly higher than theists. Agnostics occupy an intermediate position. The mean ratings for arguments for theism were 25.5 (SD = 5.7) for theists, 13.4 (SD = 5.8) for atheists and 16.6 (SD = 5.2) for agnostics. The mean ratings for arguments against theism were 25.6 (SD = 6.9) for atheists, 17.7 (SD = 4.7) for theists and 21.5 (SD = 5.8) for agnostics. A non-parametric test, the Kruskal–Wallis ANOVA, shows that these differences between theists, atheists and agnostics are statistically significant, both for arguments for theism, $df(2) = 397.2$, $p = .0001$ and for arguments against theism, $df(2) = 217$, $p = .0001$. Figures 7.1 and 7.2 provide boxplots for the total scores for theism and against theism. All pairwise differences between theists, atheists and agnostics are statistically significant at the .001 level.

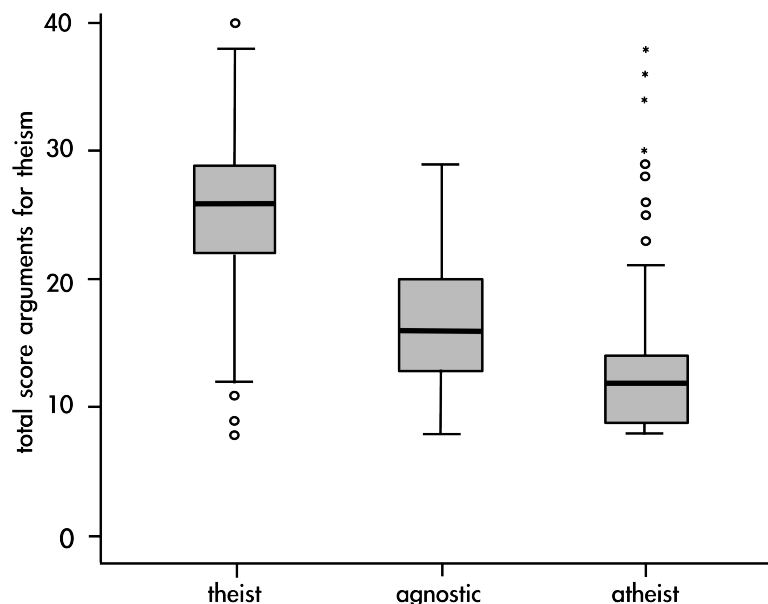


Figure 7.1: Boxplots with mean scores for all arguments for theism combined for theists, agnostics and atheists (circles and stars denote outliers).

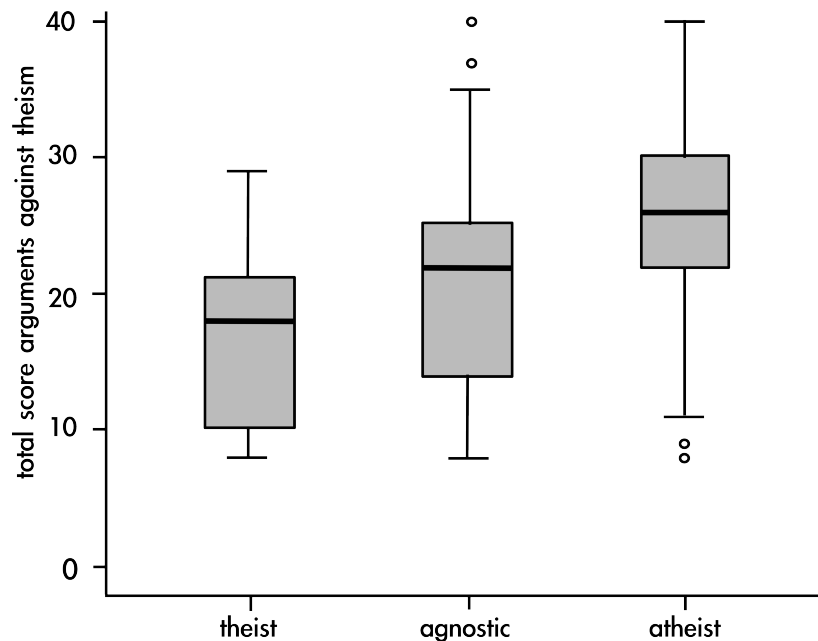


Figure 7.2. Boxplots with mean scores for all arguments against theism combined for theists, agnostics and atheists (circles denote outliers).

Looking at the arguments separately, we find significant differences between the mean ratings of each argument by theists, atheists and agnostics, with a greater difference of opinion about some arguments compared to others. We predicted that for the arguments where disagreement is largest, religious belief (atheism, theism or agnosticism) correlates more strongly with how people evaluate these arguments. Since the opinions of theists and atheists are the most divergent, we took the mean difference between their ratings of individual arguments as a quantitative measure for the disagreement. Pearson's r indicates how the religious belief of participants correlates with their ratings.

For arguments in favor of theism, the cosmological argument elicited the largest disagreement between theists and atheists, with a mean difference of 2.08 (mean theists: 3.92, mean atheists: 1.85, $r=-.483$). The moral argument shows the second largest disagreement, with a mean difference of 2.01 (mean theists: 3.40, mean atheists: 1.38, $r=-.472$). Theists and atheists disagree least about the ontological argument and the pragmatic argument for theism: the mean difference for the former is 1.29 (mean theists: 3.01, mean atheists: 1.72, $r=-.346$), the mean difference for the latter is 0.95 (mean theists: 2.59, mean atheists: 1.64, $r=-.209$). All correlations are significant at the .0001 level.

For the arguments against theism, theists and atheists disagree most strongly about the argument from lack of evidence, with a mean difference of 1.64 (mean theists: 2.16, mean atheists: 3.80, $r=.334$) and the argument from incoherence, with a mean difference of 1.58 (mean theists: 1.76, mean atheists: 3.34, $r=.389$). They disagreed least about the argument from evil and the argument from divine hiddenness. The former is the only argument where agnostics are not situated in between theists and atheists, as they rate it as less strong than either of the other groups (mean theists: 3.50, mean atheists: 3.76, mean agnostics: 3.2, mean difference between theists and atheists: .26). The correlation between religious belief and rating of the argument from evil is not significant ($r=-.013$). For the argument from divine hiddenness, the mean difference between theists and atheists is .43 (mean theists: 2.70,

mean atheists: 3.13, $r=.055$). Here also, the correlation between religious belief and rating of this argument is not significant.

Looking at how strong the arguments are rated overall, there is an interesting asymmetry between the cosmological argument (rated as the top argument for theism) and the argument from evil (rated as the top argument against theism). Whereas theists and atheists disagree strongly about the strength of the cosmological argument, and their ratings of this argument can be predicted to a significant extent from their religious beliefs, the argument from evil does not elicit the same amount of disagreement. Indeed, there is no significant correlation between the religious beliefs (atheist, theist and agnostic) of participants and their evaluation of this argument, and the difference between their mean ratings is small.

We thus found a confirmation for our prediction that religious belief significantly influences the evaluation of religious arguments: theists, atheists and agnostics differ in how they evaluate arguments for and against the existence of God. The results are highly significant overall. For each of the arguments, except for the arguments from evil and from divine hiddenness, religious belief and rating of the argument correlate significantly.

3.5 Philosophical specialization influences the appraisal of some arguments

As we have seen, religious belief strongly predicts how participants evaluate natural theological arguments. Philosophers of religion gave significantly less 'am unfamiliar with this argument' responses than non-philosophers of religion for the following arguments: the ontological argument and the arguments from divine hiddenness, inconsistent revelations, poor design and incoherence. The largest difference between specialists and non-specialists was observed for the argument from divine hiddenness: 25.8 percent of non-philosophers of religion were unfamiliar with this argument, compared to 4.4 percent of philosophers of religion ($\chi^2(1, N = 802) = 54.180, p < .0001$).

Using a logit-probit model for the arguments for theism, and a cumulative logit model without proportional odds for the arguments against theism⁴, we investigated whether philosophical specialization has an effect that is independent from religious belief. In other words, taking into account that theists are disproportionately represented in philosophy of religion, does this specialization influence how philosophers evaluate the arguments? A cautionary note before proceeding with the statistical analysis of these results: as this is an exploratory study, where we are mainly interested in detecting potential patterns and less in avoiding false positives, we have not used Bonferroni or other methods of correction for multiple comparisons⁵. Given that we performed about 50 tests, the p values would have had to be on the order of .001 or smaller to reach significance using Bonferroni correction. While the effects of religious belief are robust, with large effect sizes for correlations of religious belief and appraisal of arguments, the effects for philosophical specialization and gender are more modest.

Participants who have philosophy of religion as an area of specialization evaluated several arguments for theism more positively than those who do not. Philosophers of religion are 1.53 times more likely than non-philosophers of religion to rate the cosmological argument favorably ($p = .01$). For the argument from design, philosophers of religion are 1.76 times more likely than non-philosophers of religion to rate it favorably ($p = .0005$). Philosophers of religion are 1.42 times more likely than other philosophers to rate the argument from miracles favorably ($p = .043$). They

are also 1.39 times more likely than others to rate the argument from religious experience favorably ($p = .042$).

In this survey, the only argument against theism where philosophy of religion had a positive effect was the argument from divine hiddenness. Philosophers of religion were 1.54 times more likely than participants who do not have this area of specialization to rate this argument as 'strong' versus 'neutral,' 'weak' or 'very weak' ($p = .03$). Philosophers of religion rated several arguments against theism as weaker. Non-philosophers of religion are 1.82 times more likely than philosophers of religion to rate the argument from parsimony as 'strong' versus 'neutral,' 'weak' or 'very weak' ($p = .02$). Non-philosophers of religion are 2.26 times more likely than philosophers of religion to rate the pragmatic argument for atheism as 'strong' versus 'neutral,' 'weak' or 'very weak' ($p = .02$). Subjects who are not philosophers of religion are 1.58 times more likely than those who are to rate the argument from incoherence as 'neutral' versus 'weak' or 'very weak' ($p = .03$), and they are 1.76 times more likely than philosophers of religion to rate it as 'strong' versus 'neutral,' 'weak' or 'very weak' ($p = .03$).

3.6 Gender effects

Only 9.6 percent of the respondents who were philosophers of religion were female, compared to 24.2 percent women in the entire sample. When focusing on faculty members (tenure-track, tenured, visiting assistant professors and postdocs, $N=391$), there were 12.3 percent women in philosophy of religion. The difference between the representation of women in philosophy of religion and in other philosophical specializations is statistically significant, $X^2(1, 391) = 11.507$, $p = .001$.

In this survey, the percentage of theists among female respondents (22.2 percent) was significantly lower than the percentage of theists among male respondents (46.4 percent), $X^2(1, N=802) = 35.78$, $p = .0001$. To examine whether gender has an overall effect on the assessment of the arguments, we used a two-way ANOVA with religious belief and gender as independent variables. Controlling for religious belief, we found no difference, $F(11, 802) = 3.28$, $p = .071$ between how strongly male and female participants rated the arguments for theism combined and the arguments against theism combined.

Given the recent interest in the effects of gender on philosophical opinion (e.g., Buckwalter and Stich 2014; Seyedsayamdost in press) we also examined whether gender is significantly correlated with assessing individual arguments. Male respondents found the cosmological argument and the argument from evil stronger. Female participants found the ontological argument, the pragmatic arguments for theism and atheism, and the arguments from inconsistent revelations and incoherence stronger.

In this survey, men were 1.52 times more likely than women to rate the cosmological argument more favorably ($p = .01$). They were 1.66 times more likely to rate the argument from evil as 'strong' versus 'neutral,' 'weak' or 'very weak' ($p = .004$). Women were 1.43 times more likely than men to rate the ontological argument favorably ($p = .029$). They were 1.59 times more likely than men to rate the pragmatic argument for theism more favorably ($p = .005$). They were 1.53 times more likely than men to rate the argument from inconsistent revelations as 'neutral' versus 'weak' or 'very weak' ($p = .02$); and 1.53 times more likely to rate it as 'strong' versus 'neutral,' 'weak' or 'very weak' ($p = .03$). They were 1.81 times more likely than men to rate the argument from incoherence as 'neutral' versus 'weak' or 'very weak' ($p = .004$), and 1.71 times more likely than men to rate the argument from

incoherence as ‘very strong’ versus ‘strong,’ ‘neutral,’ ‘weak’ or ‘very weak’ ($p = .04$). They were 1.75 times more likely than men to rate the pragmatic argument for atheism as ‘weak’ versus ‘very weak’ ($p = .01$), 2.27 times more likely to rate it as ‘neutral’ versus ‘weak’ or ‘very weak’ ($p = .0001$), 1.85 times more likely than men to rate it as ‘strong’ versus ‘neutral,’ ‘weak’ or ‘very weak’ ($p = .01$), and 2.22 times more likely than male respondents to rate it as ‘very strong’ versus ‘strong,’ ‘neutral,’ ‘weak’ or ‘very weak’ ($p = .03$). Remarkably, in this sample the arguments that men find stronger are also rated stronger overall, whereas several of the arguments that women find stronger are overall rated as weak—women also rate these arguments as relatively weak, but less so than men do.

For the argument from beauty and the argument from miracles, there is a significant interaction between gender and religious belief. Male theists were 26.83 times more likely than male atheists, and 7.72 times more likely than female atheists to rate the argument from beauty more favorably (both $p < .0001$). They were 27.27 times more likely than male atheists, and 12.7 times more likely than female atheists to rate the argument from miracles more favorably (both $p < .0001$). Male theists were 10.26 times more likely than male agnostics, and 3.56 times more likely than female agnostics to rate the argument from miracles more favorably (both $p < .0001$). Female theists are 20.93 times more likely than male atheists and 9.75 times more likely than female atheists to rate the argument from miracles more favorably (both $p < .0001$).

4. Discussion

This study has found several factors that significantly influence the appraisal of natural theological arguments for or against theism. The strongest predictor is religious belief, followed by philosophical specialization and gender. We will now discuss why these factors may play a role.

4.1 Confirmation bias

Several authors (e.g., Levine 2000; Trakakis 2008; Draper and Nichols 2013) have hypothesized that the religious beliefs of philosophers of religion (who are, to a large extent, Christian theists) unduly influence their appraisal of religious arguments. More specifically, theist philosophers of religion would evaluate the arguments in a selective way so that it supports their prior views.

People tend to evaluate evidence and arguments that are in line with their beliefs more favorably, and to dismiss them when they do not support their beliefs. For instance, Taber and Lodge (2006) found that people consistently rate arguments in favor of their views on gun control and affirmative action more strongly than arguments that are incongruent with their views. When respondents could freely pick and choose information to look at, most of them actively sought out sympathetic, nonthreatening sources, e.g., those in favor of gun control were less likely to read the sources against gun control that were presented to them and vice versa. Confirmation bias is the tendency to look for evidence that confirms, rather than disconfirms, one’s earlier beliefs. It is accompanied by disconfirmation bias, the tendency to dismiss evidence or arguments for beliefs that are incompatible with one’s own convictions (see Nickerson 1998; Mercier 2010 for review). These biases are not attenuated by education. For instance, molecular biologists—especially inexperienced ones—who find anomalous data tend to dismiss them as errors or experimental artifacts, rather than as potential disconfirming evidence (Fugelsang et al. 2004). The robustness of confirmation bias in highly educated people leads to the prediction that philosophers

find arguments that support beliefs they already hold more persuasive than arguments that go against their prior beliefs.

Mere influence of prior beliefs on evaluation of arguments is indicative, but not conclusive evidence for confirmation bias. Suppose that philosophers all believe that $1+2=3$ and are given an argument with a true mathematical conclusion that uses this as a central premise. The fact that philosophers come to believe the conclusion of this argument indicates that their prior beliefs influence their evaluation of the argument, but is not evidence of confirmation bias. However, assume that some philosophers believe a controversial philosophical thesis (e.g., that free will is compatible with determinism), whereas others do not. If the ratings by philosophers of arguments for or against compatibilism are significantly influenced by their prior beliefs, it would seem that confirmation bias plays a role. This would be especially the case if few philosophers changed their minds as a result of such arguments.

Our findings are consistent with a significant role of confirmation bias in philosophy of religion. Importantly, this bias is present in both theists and atheists. Atheists rate the arguments against theism more strongly than theists, and find the arguments for theism weaker than theists, and vice versa. Although being specialized in philosophy of religion makes a difference for the assessment of some arguments, religious belief remains the best predictor for how the arguments are rated.

When the first author reported preliminary findings of this survey on the philosophy blog Prosblogion, a number of philosophers expressed disapprobation at this significant role of confirmation bias, with one commenter writing ‘My reading of the situation is that philosophy of religion is unhealthy.’ Psychologists who study confirmation bias (e.g., Chinn and Brewer 1998) mostly regard it as having a negative impact on assessment and decision-making (hence the term ‘bias’) as it leads people to ignore relevant evidence, to twist anomalous evidence to fit their expectations and to fail to consider potentially viable alternatives. On the other hand, a disconfirmation strategy is often regarded as one that leads to sounder reasoning. This slipping of normative assumptions into psychological work is common (see Elqayam and Evans 2011 for review); it often has a significant impact on how empirical studies are designed. For instance, in confirmation bias studies, participants often do not receive alternative explanations, and working memory demands are heavy, which might make confirmation the only viable strategy. Using a carefully controlled framework in which participants were asked to use either a confirming or disconfirming strategy to evaluate evidence for a given theory (e.g., why whales get beached), Koslowski et al. (2013) found that disconfirmation does not automatically lead to a more careful assessment of potential sources of evidence than a confirmation strategy. In fact, participants who had to disconfirm a given hypothesis were more likely to think that evidence that was in fact neutral to the hypothesis disconfirmed it. As Koslowski et al. (2013) indicate, the term ‘confirmation bias’ is actually a conflation of several forms of reasoning, running the gamut from a reasonable strategy (finding evidence to support a given hypothesis) to wishful thinking (deliberately ignoring potential disconfirming evidence).

Recently, proponents of the argumentative theory of reasoning (e.g., Mercier and Sperber 2011) have suggested that confirmation bias and other reasoning biases are not flaws but that they are conducive to good reasoning. They are only flaws if we see reasoning as a solitary process of a detached, Cartesian mind. In order to reach a balanced conclusion such a mind needs to be able to survey all the possibilities. However, when people are reasoning in dialogue with each other, such caution might

not be necessary, and it might be more prudent to maintain one's belief in the face of disagreement if one thinks one is right: 'In group settings, reasoning biases can become a positive force and contribute to a kind of division of cognitive labor' (Mercier and Sperber 2011, 73). Some studies (e.g., Trouche et al. 2014) indicate that people who are right are more likely to convince others in argumentative contexts than people who are more confident. In these studies, participants had to solve a puzzle with a non-obvious solution. Those who found the right answer tend to be better at convincing the others, because they have better-grounded arguments. Results like these indicate that even if individuals are biased, interaction through argumentation can lead to sounder reasoning. However, the participants in this study did not approach the puzzle with high or low prior probabilities attached to possible solutions, which is the case in the religious domain.

Even if the argumentative theory of reasoning is correct, it might be problematic that philosophers of religion evaluate the plausibility of theism using a confirmation strategy, which leads them to rate arguments in line with their prior beliefs more positively. Polarized debates in the public sphere, e.g., about climate change, vaccines and evolutionary theory indicate that the invisible hand of argumentative reasoning does not always work well, especially not in cases where one's own preconceptions play a significant role in the evaluation of arguments. This happens, for instance, if debaters are so entrenched in their views that they do not listen to their opponents, or even dismiss them out of hand as irrational (Morin 2014). Moreover, in cases where the invisible hand of argumentation works effectively—where biases at the individual level are corrected in argumentative contexts—it can only do so if there is a diversity of opinions. Thus, the large percentage of Christian theists in analytic philosophy of religion is somewhat worrisome. Among the atheist minority of philosophers of religion, the majority are naturalists. As Draper and Nichols (2013) and Schellenberg (in press) observe, it is striking that (Christian) theism and scientific naturalism are the only options that are systematically considered in philosophy of religion, to the extent that disconfirming evidence for one position is automatically seen as confirming evidence for the other: 'most naturalists too assume that theistic God-centered religion must succeed if any does. Naturalism or theism. These seem to be the only options that many see. The harshest critics of religion, including philosophers such as Daniel Dennett, seem to think their job is done when they have, to their own satisfaction, criticized personalistic, agential conceptions of a divine reality' (Schellenberg in press). The role of confirmation bias in evaluating religious arguments can be mitigated by a broader pluralism in philosophy of religion.

4.2 Early-developed intuitions

We have argued elsewhere (e.g., De Cruz and De Smedt 2010; De Smedt and De Cruz 2011; De Cruz and De Smedt 2015) that arguments in natural theology draw on intuitions and cognitive tendencies that arise developmentally early and robustly in human cognition. For example, the cosmological argument, which infers the existence of God from the existence of the universe, draws on causal intuitions we employ in everyday life. When an event occurs, we have the spontaneous intuition that something has caused it. Even infants and nonhuman animals have such causal intuitions. Experiments (e.g., Saxe et al. 2005; Newman et al. 2010) indicate that preverbal infants expect that something that happens, such as a stack of disordered blocks becoming ordered, has an external cause, and they prefer agents (e.g., a human hand) as causes over non-agents (e.g., a toy train).

The design argument relies on the intuition that ordered complexity is the result of purposive design. Features of the natural world, such as the eye or the cosmological constants and laws of nature, seem fine-tuned and designed for the functions they fulfill. A large, cross-cultural literature reveals that young children develop a strong preference for teleological explanations for the world they live in (see Kelemen 2004 for a review) and that these intuitions remain present throughout life.

The moral argument proposes that there are objective moral norms, and that the existence of such norms is more probable under theism than under naturalism. It capitalizes on our intuition that moral norms are objective, rather than purely contingent and subjective. From an early age onward, humans find moral conventions more objectively true than other norms, such as dress codes or taste preferences. For instance, young children think that hitting and pulling hair is not allowed in other countries, but they believe that dress codes may be more variable (Nichols and Folds-Bennett 2003).

The cosmological, design and moral arguments depend on intuitions about causation, teleology and ethics that are found in neurotypical members of our species; they are not specific to theism. This is in line with our empirical finding that these arguments are among the most highly rated arguments for theism. Remarkably, they also elicit a high degree of disagreement, with the cosmological and moral arguments eliciting the largest disagreement between theists and atheists.

If the premises on which these arguments depend rely on intuitions that are widespread, why do atheists and agnostics disagree about their conclusions? Religious arguments may be question-begging in a doxastic sense (Faust 2008). Although they are not strictly speaking circular (i.e., they do not contain their conclusions in their premises), reasoners accept their premises more readily if they also accept their conclusions. Thus, although it is intuitively plausible that moral norms are objective, nontheists tend to deny this premise of the moral argument and argue that moral norms are not really objective. Indeed, over the past decades, naturalists have offered detailed arguments for the claim that naturalism is incompatible with moral realism (e.g., Street 2006; Joyce 2006), and have also speculated why moral realism seems intuitively true (e.g., Ruse 2010). Similarly, although our everyday intuitions make the premise ‘whatever begins to exist has a cause for its existence’ (the first premise of the *kalām* cosmological argument) plausible, nontheists routinely deny this premise, and argue that the universe, unlike things in our everyday experience, caused itself to exist (Smith 1999).

4.3 Philosophical expertise

The existence of God is a classic case of peer disagreement where disagreeing parties are similarly educated and have access to the same body of evidence (perhaps with the exception of private evidence such as religious experience and *sensus divinitatis*), and it is therefore often discussed in the literature on peer disagreement (e.g., Frances 2014). Relevant evidence includes our intuitions about features of the world, such as (perceived) design and fine-tuning, the existence of the cosmos, moral norms, the pervasiveness of suffering among humans and other sentient beings, and the observation that God’s existence is not an obvious fact (hiddenness). Several authors argue that keeping steadfast in one’s own beliefs (e.g., van Inwagen 2010) is rational in the light of disagreements in natural theology. For instance, Rowe (1979) maintained that even in the face of particularly convincing cases of natural evil (such as a fawn dying slowly in a forest fire), the theist can still argue that this evil is not

gratuitous, and therefore, that an atheistic conclusion does not follow. By contrast, conciliationists, such as Christensen (2007) and Elga (2007), have argued that one's confidence should at least be somewhat shaken when confronted with a disagreeing peer who has reached a different conclusion following the same evidence.

The majority of philosophers of religion are theists. Their opinions about God's existence differ significantly from philosophers who do not specialize in philosophy of religion, only a minority of which believes that God exists. Should this opinion of specialists carry some weight? What if one's interlocutor is an epistemic superior? Take Jake, an atheist epistemologist, who believes that the arguments from evil and divine hiddenness support his atheism. He does have some knowledge of philosophy of religion (mainly from the time of his undergraduate training), but he has never done a thorough study of the subject, let alone written any papers in philosophy of religion. Jake knows there are philosophers who have thought longer and harder about the question of God's existence than he did. On the basis of the same body of evidence, they have reached the opposite conclusion that God exists. Moreover, Jake knows that the majority of philosophers of religion have reached this same conclusion. Should he revise his beliefs, or at least become less confident in his atheism?

In general, it seems sensible to defer to experts when one lacks the relevant expertise. If there is disagreement among experts, Goldman (2001, 97) recommends to look at the numbers and go with the majority, arguing that one can be 'fully justified trusting E_1 [an expert who defends the majority view] over E_2 [an expert who is part of the minority] if almost all other experts on the subject agree with E_1 , or if even a preponderance of the other experts agree with E_1 .' Given the high degree of consensus among philosophers of religion that God exists, the conclusion that one should at least accord weight to this consensus, or defer to it, seems to follow.

One can challenge this line of reasoning. Frances (2010) stipulates several cases in which one can blamelessly hold a philosophical position that is incompatible with what experts in the field think. For instance, the disagreement may be rooted in an undefended assumption, or a factor where expertise does not extend. Philosophers of religion might be religious believers for other reasons than the arguments, such as their upbringing or feeling at home in a religious community which color their evaluation of evidence. Also, self-selection might play a role in who decides to specialize in philosophy of religion: philosophers who believe that God exists, might be more motivated to study the field than those who do not. After all, if God exists, philosophy of religion is one of the most important areas of philosophy. But if he does not, it is merely on a par with the philosophy of other culturally widespread phenomena, like philosophy of music or philosophy of sport. Moreover, confirmation bias can explain to a large extent how philosophers evaluate natural theological arguments; this is the case for theists as well as atheists, for arguments for as well as against theism. An unpublished survey by the first author shows that rates of conversion in philosophers (since graduate school) are relatively small. Only 17.6 percent of philosophers reported a significant change in religious outlook, and for many of these, the change was not from theism to atheism or vice versa, but from one religious denomination to another. It seems that Jake can blamelessly believe that God does not exist while still holding that philosophers of religion, such as Eleonore Stump, Richard Swinburne and Alvin Plantinga, are his epistemic superiors in particular subjects in philosophy of religion.

Jake might reach the opposite conclusion: the reason that the majority of philosophers of religion believe theism is true has little or nothing to do with their

philosophical work. They already hold theism from the outset, and their theistic beliefs color their evaluation of the evidence. If anything, the fact that the majority of philosophers working in this field are theists points to a large role of irrelevant causal factors in reaching philosophical conclusions, which may further lower Jake's confidence in the soundness of arguments for theism. When authors like Draper and Nichols (2013) and Levine (2000) argue that philosophy of religion lacks vitality, they have this worry in mind.

Whether their misgivings are justified depends on whether permissivism is plausible. According to the uniqueness thesis (e.g., White 2005), there is only one rational attitude one can hold in the face of a total body of evidence, whereas permissivists (e.g., Schoenfield 2014, Kelly 2014) propose that there is some leeway, and that people can hold several positions in the light of the same evidence. There are different flavors of permissivism. Strong versions hold that one's beliefs are reasonable when they cohere with one's earlier beliefs (Schoenfield 2014), whereas moderate permissivists argue that only a certain range of beliefs is reasonable (Kelly 2014). From the perspective of uniqueness, only some philosophers are rational. For instance, according to Feldman (2007), suspending judgment (agnosticism) is the only rational option. For permissivists, philosophers can reasonably hold religious beliefs (including atheism) if they have good arguments for them; they can thus be significantly influenced by irrelevant causal factors and yet be rational. Take Jenny, who has been raised in an agnostic environment, by parents and teachers who never talked about God⁶. When she considers God's existence, her epistemic situation differs from that of other philosophers who were raised as theists or atheists, as she was exposed to neither firm religious beliefs nor disbeliefs. She lacks experiences and habits of mind that theists or atheists hold. The fact that her background differs from theirs has nothing to do with the truth or falsity of theism, and is thus epistemically irrelevant; it is a matter of chance that she happened to grow up in an agnostic environment. If theist philosophers are influenced by irrelevant causal factors, so are atheists, and even agnostics like Jenny (no one comes from a neutral background that would yield an epistemically privileged perspective). Given the importance of irrelevant influences in shaping religious views, even if all philosophers of religion subscribed to theism, this would still not mean theism is true. The high percentage of theists in philosophy of religion is therefore not an indication of the truth of theism. (For analogous reasons, the low percentage of theists in the general philosophy population is not epistemically relevant either.)

While Jake may rationally hold his atheistic beliefs in the face of the large percentage of theists in philosophy of religion, he might still learn something, not about the overall conclusion one should draw from natural theological arguments, but about their individual strengths. As we saw, philosophy of religion has an independent effect on the perceived strength of some arguments. The cosmological argument and the arguments from design, religious experience, miracles and divine hiddenness are rated as stronger by philosophers of religion than by other philosophers when controlling for religious belief. This may be a result of familiarity: non-philosophers of religion are not familiar with the recent literature. While they know the argument from miracles vaguely through Hume's (1748[2007]) dismissive discussion, they have little knowledge of recent scholarship, e.g., probabilistic versions of the argument that take into account biblical research (McGrew and McGrew 2009). Thus, although our survey did not reveal a lower self-reported knowledge of the argument from miracles by non-philosophers of religion, they are likely unfamiliar with the more recent versions. Likewise, the argument from divine

hiddenness is treated in specialist literature (the majority of which appears in monographs, edited volumes and philosophy of religion journals), with which non-philosophers of religion are unfamiliar. This was revealed in the self-reported familiarity with the argument, with only 4.4 percent of philosophers of religion being unfamiliar with the argument, versus 25.8 percent of other philosophers. Since differences in the evaluation of the arguments remain when controlling for religious belief, it seems that familiarity with the recent specialist literature has some epistemic weight.

4.4 Gender

Like most other areas of philosophy, philosophy of religion exhibits a gender imbalance. In this sample only 9.6 percent of philosophers of religion were female, whereas in the sample as a whole, 24.2 percent were women. Buckwalter and Stich (2014) propose that the underrepresentation of women in philosophy is partly caused by gender differences in philosophical intuitions. They provide several lines of purported evidence for gender differences in intuitions about philosophical scenarios, such as Gettier cases and Thomson's violinist. In this study, the cosmological argument and the argument from evil are rated as stronger by men, so it is unsurprising that they have an overall stronger rating, given the high percentage of male respondents in the sample. Some of the arguments that women find relatively stronger, e.g., the pragmatic arguments for theism and atheism, and the arguments from inconsistent revelations and incoherence, are overall rated as weaker.

Buckwalter and Stich (2014) is in part a meta-analysis, which relies among others on responses collected from researchers and instructors who reported gender differences. That part of their study may therefore be subject to selection bias—researchers who obtained null results may have been less inclined to report their findings. As it stands, several papers that looked for gender effects in philosophical intuitions (e.g., Adleberg et al. in press; Seyedsayamdost in press) failed to replicate their results. In this study, we were mainly interested in detecting potential effects rather than minimizing type 1 errors, so we did not correct for multiple tests, which may have colored our findings on gender effects. Since most gender effects we found are relatively modest, they should be replicated to establish whether they live up to scrutiny.

5. Concluding remarks

This paper presents findings of a large quantitative survey on how philosophers evaluate arguments for or against the existence of God. To our knowledge, this is one of the first studies in experimental philosophy of religion. It is also the first to provide a quantitative measure of the role of religious belief in evaluating natural theological arguments and, more tentatively, of the role of philosophical specialization and gender. We replicated Bourget and Chalmers's (2014) observation that a large majority of philosophers of religion are theists. We found a robust correlation between perceived strength of natural theological arguments and religious belief. Furthermore, we found a potential influence of philosophy of religion as an area of specialization, and of gender on the evaluation of these arguments.

To further an understanding of the ways in which philosophical work and religious belief interact, qualitative data are also needed⁷. Such data can shed light on the role of religious upbringing, and the relationship between privately held religious beliefs and publicly defended philosophical viewpoints. While we have focused on

the appraisal of natural theological arguments, philosophy of religion is a much broader field, with areas such as concepts of God and the metaphysics of persons (e.g., body-soul dualism). Within cognitive science of religion, there is an increasing body of literature on these topics, but at present, there is little awareness of this among philosophers.

References

Adleberg, T., M. Thompson, and E. Nahmias. 2015. Do men and women have different philosophical intuitions? Further data. *Philosophical Psychology* 28: 615–641.

Bourget, D. and D.J. Chalmers. 2014. What do philosophers believe? *Philosophical Studies* 170: 465–500.

Buckwalter, W. and S. Stich. 2014. Gender and philosophical intuition. In J. Knobe and S. Nichols, eds, *Experimental philosophy: Volume 2*, 307–316. Oxford: Oxford University Press.

Chinn, C.A. and W.F. Brewer. 1998. An empirical test of a taxonomy of responses to anomalous data in science. *Journal of Research in Science Teaching* 35: 623–654.

Christensen, D. 2007. Epistemology of disagreement: The good news. *Philosophical Review* 116: 187–217.

Dawkins, R. 2006. *The God delusion*. Boston: Houghton Mifflin.

De Cruz, H. 2014a. The enduring appeal of natural theological arguments. *Philosophy Compass* 9: 145–153.

— 2014b. Cognitive science of religion and the study of theological concepts. *Topoi. An International Review of Philosophy* 33: 487–497.

De Cruz, H. and J. De Smedt. 2010. Paley's iPod. The cognitive basis of the design argument within natural theology. *Zygon: Journal of Religion and Science* 45: 665–684.

— 2015. *A natural history of natural theology. The cognitive science of theology and philosophy of religion*. Cambridge, MA: MIT Press.

De Smedt, J. and H. De Cruz. 2011. The cognitive appeal of the cosmological argument. *Method and Theory in the Study of Religion* 23: 103–122.

Draper, P. and R. Nichols. 2013. Diagnosing bias in philosophy of religion. *The Monist* 96: 420–446.

Elga, A. 2007. Reflection and disagreement. *Noûs* 41: 478–502.

Elqayam, S. and J.S.B. Evans. 2011. Subtracting “ought” from “is”: Descriptivism versus normativism in the study of the human thinking. *Behavioral and Brain Sciences* 34: 233–248.

Faust, J. 2008. Can religious arguments persuade? *International Journal for the Philosophy of Religion* 63: 71–86.

Feldman, R. 2007. Reasonable religious disagreements. In L. Anthony, ed., *Philosophers without gods*, 194–214. Oxford: Oxford University Press.

Frances, B. 2010. The reflective epistemic renegade. *Philosophy and Phenomenological Research* 81: 419–463.

Frances, B. 2014. *Disagreement*. Cambridge: Polity Press.

Fugelsang, J.A., C.B. Stein, A.E. Green, and K.N. Dunbar. 2004. Theory and data interactions of the scientific mind: Evidence from the molecular and the cognitive laboratory. *Canadian Journal of Experimental Psychology* 58: 86–95.

Goldman, A.I. 2001. Experts: Which ones should you trust? *Philosophy and Phenomenological Research* 63: 85–110.

Gross, N. and S. Simmons. 2009. The religiosity of American college and university professors. *Sociology of Religion* 70: 101–129.

Hume, D. 1748 [2007]. *An enquiry concerning human understanding* (P. Millican, ed). Oxford: Oxford University Press.

Joyce, R. 2006. *The evolution of morality*. Cambridge, MA: MIT Press.

Kelemen, D. 2004. Are children “intuitive theists”? Reasoning about purpose and design in nature. *Psychological Science* 15: 295–301.

Keller, T. 2008. *The reason for God. Belief in an age of scepticism*. London: Hodder & Stoughton.

Kelly, T. 2014. Evidence can be permissive. In M. Steup, J. Turri, and E. Sosa, eds, *Contemporary debates in epistemology*, 298–312. Oxford: Oxford University Press.

Koslowski, B., J. Marasia, F. Vermeulen, and V. Hendrix. 2013. A disconfirming strategy is not necessarily better than a confirming strategy. *American Journal of Psychology* 126: 335–354.

Levine, M.P. 2000. Contemporary Christian analytic philosophy of religion: Biblical fundamentalism, terrible solutions to a horrible problem, and hearing God. *International Journal for Philosophy of Religion* 48: 89–119.

McGrew, T. and L. McGrew. 2009. The argument from miracles: A cumulative case for the resurrection of Jesus of Nazareth. In W.L. Craig and J.P. Moreland, eds, *The Blackwell companion to natural theology*, 593–662. Chichester: Wiley-Blackwell.

Mercier, H. 2010. The social origins of folk epistemology. *Review of Philosophy and Psychology* 1: 499–514.

Mercier, H. and D. Sperber. 2011. Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences* 34: 57–74.

Morin, O. 2014. The virtues of ingenuity: Reasoning and arguing without bias. *Topoi. An International Review of Philosophy* 33: 499–512.

Newman, G., F. Keil, V. Kuhlmeier, and K. Wynn. 2010. Early understandings of the link between agents and order. *Proceedings of the National Academy of Sciences USA* 107: 17140–17145.

Nichols, S. and T. Folds-Bennett. 2003. Are children moral objectivists? Children's judgments about moral and response-dependent properties. *Cognition* 90: B23–B32.

Nickerson R.S. 1998. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology* 2: 175–220.

Rowe, W.L. 1979. The problem of evil and some varieties of atheism. *American Philosophical Quarterly* 16: 335–341.

Ruse, M. 2010. The biological sciences can act as a ground for ethics. In F.J. Ayala and R. Arp, eds, *Contemporary debates in philosophy of biology*, 297–315. Chichester: Wiley-Blackwell.

Saxe, R., J. Tenenbaum, and S. Carey. 2005. Secret agents: Inferences about hidden causes by 10- and 12-month-old infants. *Psychological Science* 16: 995–1001.

Schellenberg, J.L. in press. Divine hiddenness and human philosophy. In A. Green and E. Stump, eds, *Hidden divinity and religious belief*. Cambridge: Cambridge University Press.

Schoenfield, M. 2014. Permission to believe: Why permissivism is true and what it tells us about irrelevant influences on belief. *Noûs* 48: 193–218.

Seyedsayamdost, H. 2015. On gender and philosophical intuition: Failure of replication and other negative results. *Philosophical Psychology* 28: 642–673.

Smith, Q. 1999. The reason the universe exists is that it caused itself to exist. *Philosophy* 74: 579–586.

Street, S. 2006. A Darwinian dilemma for realist theories of value. *Philosophical Studies* 127: 109–166.

Taber, C.S. and M. Lodge. 2006. Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science* 50: 755–769.

Trakakis, N. 2008. *The end of philosophy of religion*. New York: Continuum.

Trouche, E., E. Sander, and H. Mercier. 2014. Arguments, more than confidence, explain the good performance of reasoning groups. *Journal of Experimental Psychology: General* 143: 1958–1971.

Van Inwagen, P. 2010. We're right, they're wrong. In R. Feldman and T. Warfield, eds, *Disagreement*, 10–28. Oxford: Oxford University Press.

White, R. 2005. Epistemic permissiveness. *Philosophical Perspectives* 19: 445–459.

¹ Bourget and Chalmers (2014) is the only other quantitative survey to date that probes the religious views of professional philosophers.

² Among philosophers it is common to have more than one area of specialization.

³ There was moderate to high internal consistency in the answers of respondents: Chronbach's alpha was .886 for arguments for theism, and .749 for arguments against theism. Standard deviations of the responses to individual arguments were quite similar, ranging between 1.2 (argument from miracles) to 1.4 (cosmological argument).

⁴ We thank Robert O'Brien for calculating these statistics.

⁵ The Bonferroni correction is in many cases too conservative: it minimizes the risk of type 1 (false positive) errors, but as a result may miss many significant results. Note that even using this conservative measure, the effect of religious belief on appraisal of the arguments remains present.

⁶ We do not mean to suggest that the agnostics in our sample come from a background like Jenny's. We provide this example as a supposedly ideal epistemic situation for questions about the existence of God.

⁷ One of us is currently working on a paper that addresses this question, using a qualitative survey