A cross-cultural investigation of mind-body dualism

Suzanne Hoogeveen[1] & Michiel van Elk[2]

[1] University of Amsterdam

[2] Leiden University

Author Note

Analysis code will be provided at https://osf.io/....

Correspondence concerning this article should be addressed to Suzanne Hoogeveen, Department of Psychology, University of Amsterdam, Room 2.34, Nieuwe Achtergracht 129, 1018 WS Amsterdam, The Netherlands. E-mail: suzanne.j.hoogeveen@gmail.com

Abstract

Great study! Beautiful results.

Keywords: Religion, Mind-body, Dualism, Cross-cultural, Afterlife beliefs

Word count: X

A cross-cultural investigation of mind-body dualism

Methods

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study.

Participants

In total, 10,535 participants completed the online experiment. Of these, 340 participants (3.23%) were excluded because they failed the attention check, leaving an analytic sample of $N = 10195$ from 24 countries (see Table 1 for descriptive statistics per country). Participants were recruited from university student samples, from personal networks, and from representative samples accessed by panel agencies and online platforms (MTurk, Kieskompas, Sojump, TurkPrime, Lancers, Qualtrics panels, Crowdpanel, and Prolific). Participants were compensated for participation by a financial remuneration, the possibility for a reward through a raffle, course credits, or no compensation. There were no a priori exclusion criteria; everyone over 18 years old could participate. Participants were forced to answer all multiple choice questions, hence there was no missing data. The countries were convenience-sampled (i.e., through personal networks), but were selected to cover all 6 continents and include different ethnic majorities and religious majorities (Christian, Muslim, Hindu, Jewish, Eastern religions, as well as highly secular societies). Table 1 displays the method of recruitment and compensation per country.

Sampling Plan. We preregistered a target sample size of $n = 400$ per country and 20-25 target countries.The preregistered sample size and composition allowed us to look at overall effects, effects within countries, and between countries. As we applied a Bayesian statistical framework, we needed a minimum of 20 countries to have sufficient data for accurate estimation in cross-country comparisons (Hox, van de Schoot, & Matthijsse, 2012). However, our main interest were overall effects - rather than effects for individual countries.

With approximately 8,800 participants, we would have sufficient data to reliably estimate overall effects, especially since the dualism effect (state contrast) is within-subjects. As a stopping rule, we preregistered that data collection would be terminated (a) when the target of $n = 400$ per country was reached, or (b) by September 30th, 2019. However, due to unforeseen delays in construction of the materials and recruitment, this deadline was extended until November 30th, 2019.[1] We did not download or inspect the data until after November 30th. Finally, we preregistered to only include countries where usable data from at least 300 participants were collected. However, we decided to keep the $n = 291$ participants from Lithuania in the final sample, as the hierarchical models account for uncertainty in estimates from countries with smaller samples and removing these data will actually reduce the overall precision of the estimates. Moreover, it would simply be unfortunate to remove all data from an entire –highly understudied– country.

Material

The study was part of a larger project on cross-cultural effects related to religiosity (see Appendix for details about the project). The relevant variables for the current study were individual religiosity, target state category (mental state vs. physical state), the manipulated context of the narrative and the binary continuity judgments for each state. Participant religiosity was measured using standardized items taken from the World Values Survey [WVS; World Values Survey (2010)], covering religious behaviours (institutionalized such as church attendance and private such as prayer/mediation), beliefs, identification, values, and denomination. Besides having high face-validity, these measures have been applied cross-culturally in other studies (Lindeman, Svedholm-Hakkinen, & Lipsanen, 2015; Lun & Bond, 2013; Stavrova, 2015). A Bayesian reliability analysis using the `Bayesrel` package (Pfadt & van den Bergh, 2020) indicated good internal consistency of the religiosity measure, McDonald omega = 0.930 [0.927, 0.931] (all item-rest correlations > 0.61). All

[1] Ten participants completed the survey after this termination date, as the survey links were kept open (though no longer disseminated).

individual religiosity items were transformed on a 0-1 scale (to make each item contribute equally to the scale), tallied to create a religiosity score per participant, and grand-mean standardized for the analyses. The experimental stimuli consisted of a short narrative about a young person whose grandmother dies. The context was manipulated (between-subjects) by either introducing a priest (or other religious authority) or a doctor to mention the grandmother's death and stating that she is either *with God now* or *death now*, respectively. Participants then indicated whether they thought that the grandmother was still capable of (1) *being hungry*, (2) *hearing voices*, still had (3) *a functioning brain* , could still (4) *know things*, (5) *love*, and (6) *want things*. The first three processes were classified as physical states (psychobiological/perceptual) and the last three as mental states (cognitive). The narratives and process items were based on the materials used by Harris and Giménez (2005). The name of the target person and the specific religious authority were adjusted to the language and cultural context of each country.

Procedure

Participants received a link to the Qualtrics survey, either by email, social media or through an online platform. After reading the instructions and providing informed consent, they first completed items for a separate study about religiosity and trustworthiness and religiosity and source credibility (see Hoogeveen et al., 2022). Subsequently, they were presented with the short narrative in either the religious or secular context, provided continuity judgments for the six process items, and completed the manipulation check to validate that they registered the type of authority (religious vs. medical). Finally, they provided demographics, a quality of life scale, the religiosity items, and were given the opportunity to provide comments. It took about 10 minutes to complete the entire survey (median completion time was 11.4 minutes).

Data analysis

Analyses were carried out in R[2]. The models were built using the package `brms`(Bürkner, 2017b), which relies on the Stan language (Carpenter et al., 2017). The `bridgesampling` package (Gronau, Singmann, & Wagenmakers, 2020b) was used to estimate the log marginal likelihood of the models of interest and calculate Bayes factors.

The multilevel Bayesian modeling approach allows us to systematically evaluate the evidence in the data under different models: (i) across all countries the effect is truly null; (ii) all countries share a common nonzero effect; (iii) countries differ, but all effects are in the same (predicted) direction; and (iv) in some countries the effect is positive whereas in others the effect is negative (Haaf & Rouder, 2017; Rouder, Haaf, Davis-Stober, & Hilgard, 2019). The models differ in the extent to which they constrain their predictions, from the most constrained (i) to completely unconstrained (iv). We refer to these models as the null model, the common effect model, the positive effects model, and the unconstrained model, respectively. Note that while the predictions from model (iii) are less constrained than those from model (ii), it is more difficult to obtain evidence for small effects under the latter model because it assumes that the effect is present in every country, rather than only in the aggregate sample. When applied to our hypothesis for the mental vs. physical state effect, evidence for (i) would indicate that people from these 24 countries do not differentially evaluate continuity of physical and mental states after death, evidence for (ii) would indicate

[2] For all analyses, we used R [Version 4.1.2; R Core Team (2021)] and the R-packages *BayesFactor* [Version 0.9.12.4.3; Morey and Rouder (2021)], *bridgesampling* [Version 1.1.2; Gronau, Singmann, and Wagenmakers (2020a)], *brms* [Version 2.16.3; Bürkner (2017a); Bürkner (2018); Bürkner (2021)], *cmdstanr* [Version 0.4.0; Gabry and Češnovar (2021)], *coda* [Version 0.19.4; Plummer, Best, Cowles, and Vines (2006)], *cowplot* [Version 1.1.1; Wilke (2020)], *dplyr* [Version 1.0.7; Wickham, François, Henry, and Müller (2021)], *ggdist* [Version 3.0.1; Kay (2021)], *ggmosaic* [Version 0.3.3; Jeppson, Hofmann, and Cook (2021)], *ggplot2* [Version 3.3.5; Wickham (2016)], *ggsci* [Version 2.9; Xiao (2018)], *gridExtra* [Version 2.3; Auguie (2017)], *invgamma* [Version 1.1; Kahle and Stamey (2017)], *kableExtra* [Version 1.3.4; Zhu (2021)], *MASS* [Version 7.3.54; Venables and Ripley (2002)], *Matrix* [Version 1.3.4; Bates and Maechler (2021)], *MCMCpack* [Version 1.6.0; Martin, Quinn, and Park (2011)], *papaja* [Version 0.1.0.9997; Aust and Barth (2020)], *posterior* [Version 1.2.0; Vehtari, Gelman, Simpson, Carpenter, and Bürkner (2021a)], *Rcpp* [Version 1.0.8; Eddelbuettel and François (2011); Eddelbuettel and Balamuta (2018)], *rstan* [Version 2.21.3; Stan Development Team (2021)], *scales* [Version 1.1.1; Wickham and Seidel (2020)], *StanHeaders* [Version 2.21.0.7; Stan Development Team (2020)], *tidybayes* [Version 3.0.2; Kay (2022)], and *tidyr* [Version 1.1.4; Wickham (2021)].

that on average people from these 24 countries consider mental states more likely to continue than physical states (or vice versa), evidence for (iii) would indicate that in all of the 24 countries, people consider mental states more likely to continue than physical states (or vice versa), but there is cultural variation in the size of this effect, and evidence for (iv) would indicate that in some countries people consider mental states more likely to continue than physical states, and in other countries people consider physical states more likely to continue than mental states, indicating cultural variation in the direction (and size) of the effect. For the full model including all main effects and relevant interaction effects, we specified the following unconstrained model. Let Y denote the continuity responses per participant aggretaged over the three binary items per state, where 0 indicates discontinuity and 1 indicates continuity and $Y = 0, \ldots, 3$. Further, let Y_{ijkl} be the continuity judgment for the ith participant, $i = 1, \ldots, N$, in the jth country, $j = 1, \ldots, 24$, for the kth state category, $k = 1, 2$ (physical or mental states, respectively), and the lth context condition, $l = 1, 2$ (secular or religious context, respectively). The responses Y_{ijkl} are modeled using an aggregated binomail model with a logit link to transform probabilities into real numbers $\in (-\infty, \infty)$:

$$Y_{ijkl} \overset{ind}{\sim} \text{Binomial}(3, p_{ijkl}),$$

$$\text{logit}(p_{ijkl}) = \alpha_j + x_k\beta_j + u_i\delta_j + c_l\gamma_j + v_{ki}\theta_j + w_{kl}\zeta_j.$$

where $\text{logit}(p_{ijkl})$ is the combined effect of observations, countries, and state categories on the tendency to indicate 'continues.' Note that $\text{logit}(p_{ijkl}) = 0$ reflects a probability of 0.5 of indicating continuity. The term α_j serves as the baseline continuity intercept for the jth country. The indicator $x_k = -0.5, 0.5$ if $k = 1, 2$, respectively, where $k = 1$ indicates the physical state condition and $k = 2$ indicates the mental state condition. The term β_j is the jth country's main effect of state category on continuity judgments. The variable u_i gives the ith participant's standardized religiosity score and δ_j is the jth country's main effect of religiosity. The indicator $c_l = -0.5, 0.5$ if $l = 1, 2$, respectively, where $l = 1$ indicates the

secular context condition and $l = 2$ indicates the religious context condition. The term γ_j is then the jth country's main effect of context. The indicator v_{ki} gives the state-by-religiosity interaction term and θ_j is the corresponding interaction effect for the jth country. Finally, indicator w_{kl} gives the state-by-context interaction term and ζ_j is the corresponding interaction effect for the jth country.

Results

```
## [1] "1-to-17.96" "1-to-1"     "1-to-2.96"  "1-to-10.10"
```

```
## [1] "1-to-334"  "1-to-383"  "1-to-1695" "1-to-1"
```

On average, people made continuity judgments for 31.76% of the states, with 16.13% for physical states and 47.39% for mental states. In Figure 3A these observed rates are further unpacked per context condition and level of religiosity. Additionally, 39.05% of participants judged none of the states to continue after death, while 2.01% reported all six states to continue.

As can be seen in Table 3, we found substantial evidence in favor of our hypotheses for the state effect ($\mathcal{H}_1$), the religiosity effect ($\mathcal{H}_2$), the context effect ($\mathcal{H}_3$), and the state-by-religiosity interaction effect ($\mathcal{H}_4$), yet strong evidence against the state-by-context effect ($\mathcal{H}_5$).

First, mental processes are judged as more likely to continue after death than psychobiological processes, to a varying degree across countries: $\text{BF}_{+0} = \infty$; $\text{BF}_{+1} = 10^{26}$, $\mu_\beta = 1.71$ [1.55, 1.86], $\sigma_\beta = 0.35$ [0.25, 0.50]. This effect translates into an increase of 0.328 [0.128, 0.512] on the probability scale. Second, religiosity is positively associated with continuity judgments, to a varying degree across countries: $\text{BF}_{+0} = \infty$; $\text{BF}_{+1} = 10^{87}$, $\mu_\delta = 0.84$ [0.71, 0.96], $\sigma_\delta = 0.28$ [0.21, 0.39]. In other words, the most religious participants are 46.8% [13.5%, 71.9%] more likely to make continuity judgments than the least religious

participants. Third, people are more likely to make continuity judgments in a religious context than in a secular (medical) context, to a varying degree across countries: $BF_{+0} = 10^{146}$; $BF_{+1} = 10^{11}$, $\mu_\gamma = 0.52$ [0.41, 0.61], $\sigma_\gamma = 0.22$ [0.15, 0.32]. That is, people are 9.8% [0.9%, 21.2%] more likely to make continuity judgments in the religious context than in the secular context. Fourth, the difference in continuity judgments between mental and physical states becomes larger with increased religiosity, to a varying degree across countries: $BF_{10} = 10^{16}$; $BF_{+1} = 3126$, $\mu_\theta = 0.24$ [0.14, 0.33], $\sigma_\theta = 0.18$ [0.11, 0.28]. That is, overall, the most religious participants make an estimated 43.4% [22.9%, 57.6%] more continuity judgments about mental processes than about physical processes, while this difference is only 17.4% [3.9%, 41.4%] for the least religious participants. Note, however, that while the model comparison indicated substantial evidence for the interaction effect, the unconstrained model slightly outperforms the positive-effects model: $BF_{u+} = 1.19$. This is due to the fact that when looking at the countries separately, for 7 of them, the credible interval of the interaction effect does include zero (see Figure XXX). Fifth, the difference in continuity judgments between mental and physical states is not larger in a religious than in a secular context: $BF_{01} = 40.34$, $\mu_\zeta = -0.09$ [-0.19, 0.00], $\sigma_\zeta = 0.08$ [0.00, 0.22].

At the same time, the modal response across most countries is complete cessation rather than continuity: in only 5 out of 24 countries, the modal sum score across the six items was either 3 or 4, in all other countries it was 0. Specifically, in China, India, Japan, Romania, Singapore participants were more likely to indicate continuity of some states than complete cessation. Across the aggregated sample, the mode is also complete cessation.

Country-level cultural norms. Correlation between country-level intercepts and state effects with aggregated cultural norms of religion?

Mirroring the religiosity effect at the individual level, we expected a *positive* relation between both the overall continuity judgments and cultural norms of religion and between mind-body dualism operationalized as the state effect and cultural norms of religion. To

investigate this effect, we correlated cultural norms aggregated at the country-level with country-level estimates of the intercepts (α_j) and state-effects (β_j) in the models. First, we find some anecdotal evidence against a positive correlation between the country-level overall probability of continuity and cultural norms of religion: $BF_{+0} = 0.32$; $BF_{0+} = 3.09$. Second, we also obtained moderate evidence against a positive correlation between country-level estimates of dualism (i.e., the state effect) and cultural norms of religiosity aggregated at the country-level: $BF_{+0} = 0.13$; $BF_{0+} = 7.66$ (see Figure 7). In fact, if anything, the correlation appears to be negative, rather than positive; the estimated size of the correlation coefficient is -0.48 [-0.72, -0.16].[3]. This suggests that participants from countries where religion is more normative are *not* more likely to make continuity judgments or reason dualistically. Instead, in more religious countries, people may be less likely to distinguish between physical and mental states.

Correlation explicit afterlife beliefs. Afterlife beliefs as a predictor of intercepts and state effects (individual level). To test the hypothesis that explicit afterlife beliefs are related to both overall continuity judgments and mind-body dualism, we constructed the models used to test H2 with the item on afterlife beliefs as the predictor.

The Bayes factor analysis provided strong evidence that explicit afterlife beliefs are positively related to both the overall probability of making continuity judgments, to a varying degree across countries ($BF_{+0} = \infty$; $BF_{+1} = 10^{96}$, $\mu_\delta = 0.90$ [0.78, 1.01], $\sigma_\delta = 0.28$ [0.21, 0.39]). In addition, afterlife beliefs were also related to the tendency to differentiate between mental and physical states, to a varying degree across countries ($BF_{+0} = 10^{12}$; $BF_{+1} = 2.99$, $\mu_\theta = 0.19$ [0.13, 0.26], $\sigma_\theta = 0.10$ [0.03, 0.18]).

So yes, even atheist extinctivists make continuity judgments, and if they do it's mostly for mental states. But at the same time, the proportion of people that display these implicit afterlife beliefs in the absence of explicit afterlife beliefs is much smaller than the proportion

[3] If we release the directional constraint, we get strong evidence in favor of a correlation: $BF_{10} = 15.69$.

of people endorsing explicit afterlife beliefs but implicitly rejecting continuity in an afterlife. That is, most people in the sample indicate to at least somewhat believe in an afterlife and rate at least one state to continue in the narrative task (no%). About 20% of participants both explicitly and implicitly reject the possibility of an afterlife (no%). Then there are no% who explicitly state to somewhat believe in an afterlife, but implicitly reject continuity of any states. Yet only yes% of participants explicitly reject an afterlife but implicitly allow for states to continue after death. This suggests that explicit afterlife beliefs may be more common than implicit ones, which seems problematic for intuitive dualism accounts that assume that especially implicit afterlife beliefs are prevalent and widespread.

Pictorial dualism item. Pictorial item as predictor of intercepts and state effect (individual level).

The pictorial dualism item was taken from Forstmann, Burgmer, and Mussweiler (2012), which was adjusted from the self-other inclusion scale by Aron, Aron, and Smollan (1992). While we find evidence that dualism as measured with the pictorial item is related to overall continuity ratings, $BF_{10} = 10^{56}$, we do not find a robust pattern across all countries. That is, we get most evidence in favor of the unconstrained model: $BF_{u1} = 10^{30}$, $\mu_\delta = 0.13$ [0.07, 0.19], $\sigma_\delta = 0.14$ [0.10, 0.20] and in 8 out of the 24 countries, the credible interval for the pictorial item as predictor includes zero. Additionally, the Bayes factor model comparison gave evidence against the pictorial item predicting mind-body dualism operationalized as the difference in continuity between mental and physical states: $BF_{10} = 0.07$; $BF_{01} = 13.76$, $\mu_\theta = 0.02$ [-0.03, 0.07], $\sigma_\theta = 0.06$ [0.00, 0.12]).

Robustness checks. In the preregistration, we specified half-cauchy priors on the standard deviation. However, the prior predictive checks showed that the fat tails of the cauchy distribution resulted in implausible predictions on the probability scale (see the Appendix for details). Following recommendations by McElreath (2016) and Betancourt, Vehtari, and Gelman (2015), we used the half-normal(0,1) prior on the country-level standard deviation instead. This resulted in more reasonable prior predictions (see Appendix).

1. hearing as a mental state
2. manipulation check failures excluded
3. level of education as covariate
4. Lithuania excluded (as preregistered)
5. cauchy priors on the between-country variance (as preregistered)

Exploratory Results

Atheist extinctivists. In addition to these overall effects, we can also look specifically at self-reported atheists who explicitly state "not at all" to believe in life after death ($n = 1513$). Among this group, people made continuity judgments for 7.39% of the states, with 2.95% for physical states and 11.83% for mental states. Additionally, 83.08% of atheist extinctivists judged none of the states to continue after death, while 0.79% reported all six states to continue.

As expected, for atheist extinctivists, the estimated intercept is considerably lower than for the overall sample: -2.92 (5.11%) vs. -0.96 (27.67%), respectively, as is the effect of mental versus physical state: 1.63 (i.e., an increase of 8.1% on the probability scale) vs. 1.71 (i.e., an increase of 32.7% on the probability scale) and the effect of religious vs. secular context: 0.78 (i.e., an increase of 3.2% on the probability scale) vs. 0.52 (i.e., an increase of 9.8% on the probability scale). This is also displayed in the Figure 8. Do note that the credible intervals for the estimates are quite wide for some countries where few people identify as atheists and deny an afterlife (e.g., India and Singapore). In general, the same pattern of results is observed for the atheist extinctivists as for the overall sample; the Bayes factor model comparison indicates most evidence for a varying positive effect of *state* ($BF_{+0} = 10^{67}$; $BF_{+1} = 16.98$) and of *context* ($BF_{+0} = 10^{15}$; $BF_{+1} = 29.11$). Again, there is no evidence that the religious context manipulation results in relatively stronger continuity judgments for mental states compared to physical states (i.e., state-by-context interaction; $BF_{01} = 7.67$; $BF_{0+} = 37.83$).

Exploratory factor analysis.

Additional Analyses. We explored whether the data provide evidence for an interaction between religiosity and context, such that the religious framing effect on continuity judgments is enhanced for religious participants in particular. The Bayes factor model comparison provided most evidence for the unconstrained model: $BF_{u0} = 334$. As shown in Figure 6a, the unconstrained model is favored because in some countries (UK, Romania) the effect is positive, whereas in others it is negative (China). However, in only three countries (UK, Romania, China) do the credible intervals exclude zero. Overall, there is no evidence in favor of a religiosity-by-framing interaction effect assuming that the framing effect on continuity is larger for religious participants: $BF_{10} = 0.87$; $BF_{01} = 1.15$ (this counts as basically no evidence either way).

Finally, we tested the evidence for a three-way interaction between state, religiosity and framing, such that mind-body dualism increases with religiosity, and particularly when framed in religious terms. The Bayes factor model comparison indicated some evidence in favor of a common three-way interaction: $BF_{10} = 17.96$. However, as shown in Figure 3b, it appears that compared to the secular framing, the religious framing slightly increases mind-body dualism for low religiosity in particular, but not for high religiosity. Based on the unclear pattern, the relatively small Bayes factor given the amount of data, and the fact that the three-way interaction only appears in 3 out of the 24 countries (see Figure 6b), we do not consider this effect of relevance.

Discussion

Prior predictive checks

In order to systematically and thoroughly assess the adequacy of the priors, we should look at some settings for both the priors on the intercepts, the effects, the variability between countries and the correlation. We can use previous studies to inform our options.

We will consider the following:

- intercept:

1. normal(0,1)
2. normal(0,5)

- effect:

1. normal(0,5)
2. normal(0,1)
3. normal(0,0.5)

- standard deviation between countries:

1. exponential(1): as suggested by McElreath in Statistical Rethinking
2. inverse-gamma(3,0.5): assuming a standard deviation below 0.5
3. cauchy(0,2): as preregistered, no idea why...
4. normal(0,1)
5. student-t(3,0,2.5): brms default

- correlation matrix:

1. lkj(1): flat distribution for the correlation matrix
2. lkj(2): putting slightly less mass on extreme correlation values (i.e., -1 and 1)

What do we know?

Based on data from previous studies that have been conducted across different cultures, we can get an idea of the expected intercepts and size of the effects. The mean state effect –the difference in the probability of continuity responses for mental vs. bodily states– across these 12 sites, taken from 4 previous studies is 0.16, so 16% with a standard deviation of 0.16

(15.70%). For the 10 context effects in the previous studies, the mean difference between a theistic/spiritual prime and the neutral/control condition is 0.10 (10.20%) with a standard deviation of 0.17 (17.50%). Based on these data, we would expect experimental effects of about 10-20% and a standard deviation between studies/countries of about 15-20%.

What do we want?

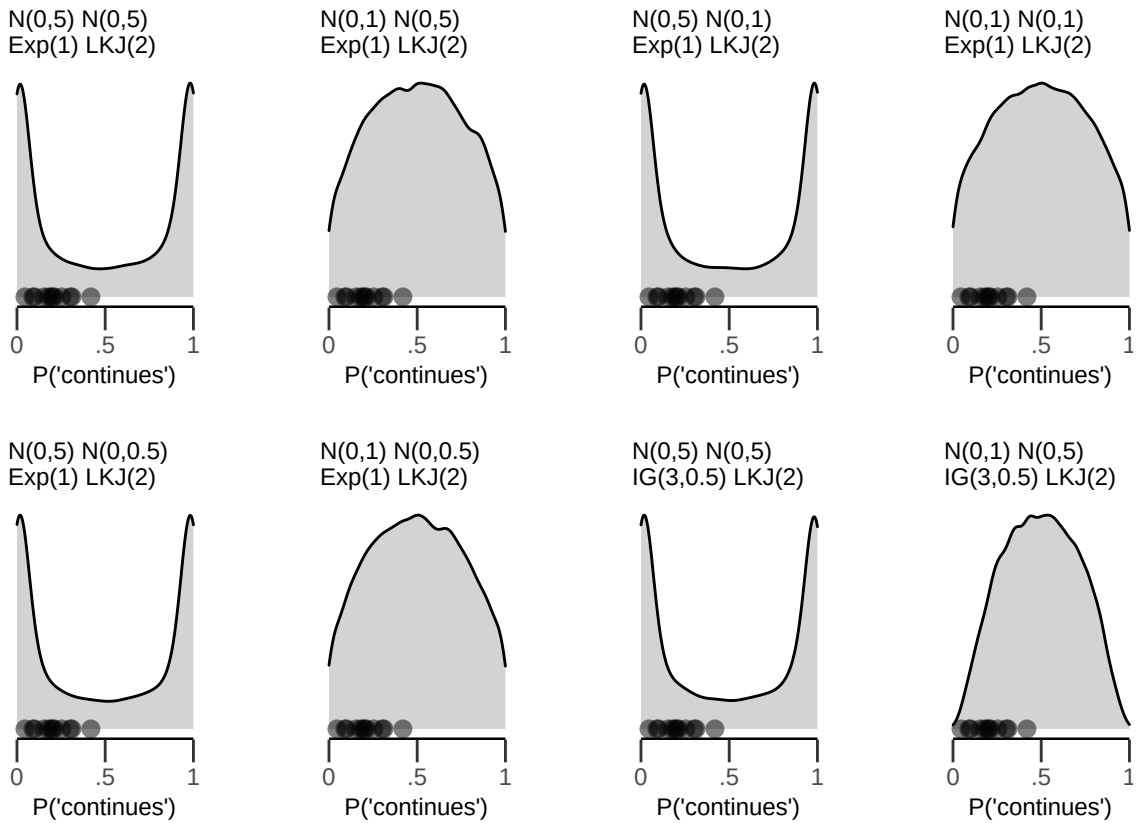
In the simulation, we draw samples from the prior distributions and look whether the distributions of the country-level intercept (i.e., the overall probability of saying that a given state will continue) and the predictions on the country-level experimental state effect (i.e., difference in probability of saying 'continues' between mental and physical states) make sense. If priors are too vague the distributions become bimodal, suggesting that all participants in a given country either judge all states to cease or continue. We aim to find prior distributions that are relatively uninformative while still allowing making sensible predictions.

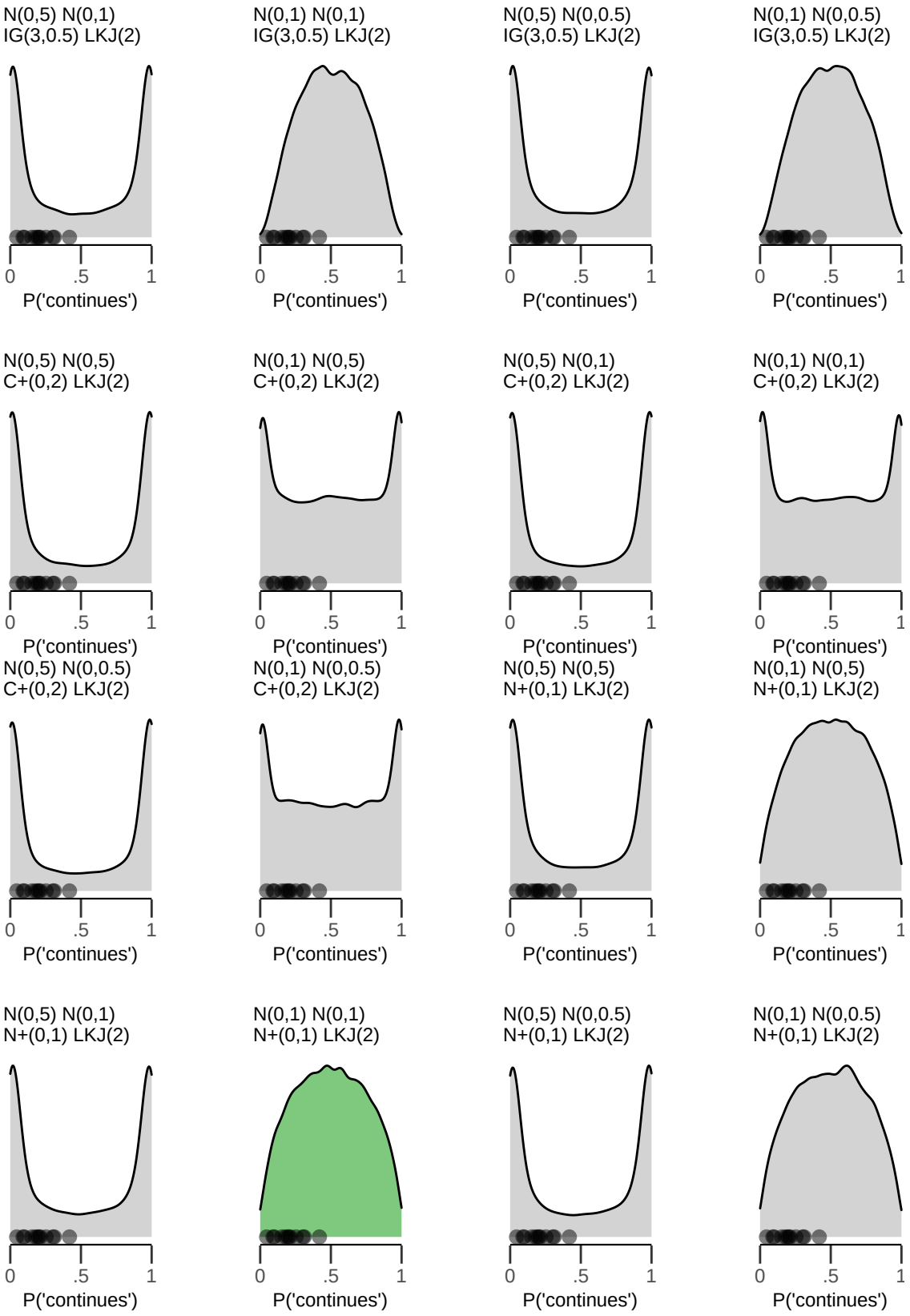
What do we conclude?

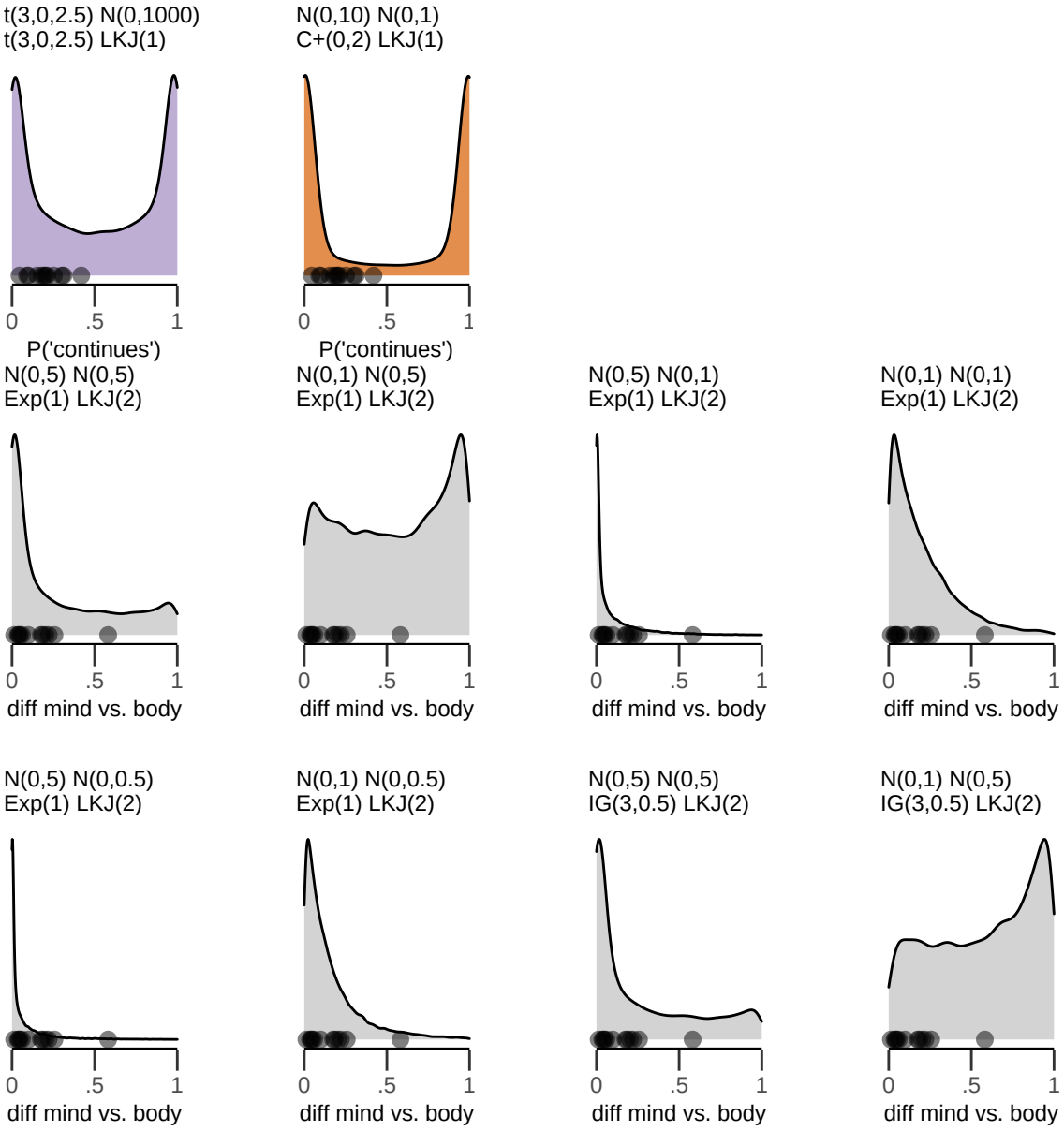
We found that the LKJ settings do not have a strong influence on the chosen parameters. We therefore show only the LKJ(2) parameter case, as we think correlations between country-level effects of -1 or 1 are less likely than more modest correlation values a priori.

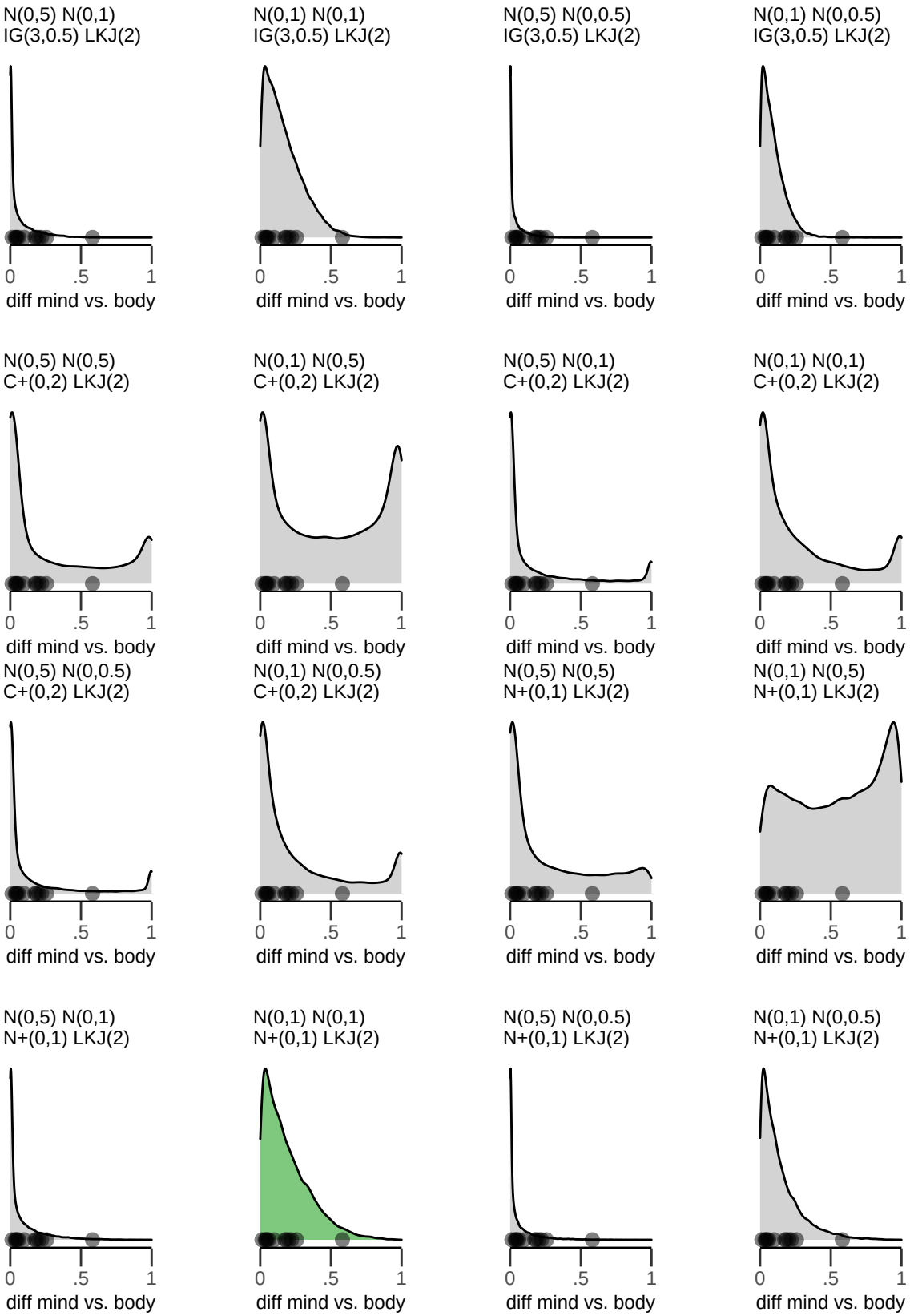
First, the normal(0,5) prior on the intercept translates into extreme predictions on the probability scale, resulting in a unrealistic bimodal distribution with most mass close to 0 and 1. The normal distribution with standard deviation 1, on the other hand, seems to make reasonable predictions about the overall probability of continuity, allowing for all values between 0 and 1 with most mass around 0.5. Second, based on visual inspection, is seems both the exponential(1) and the half-normal(0,1) prior for the between-country variation make sensible predictions. The inverse-gamma(3,0.5) seems a bit too strict and the preregistered cauchy(0,2) and the brms-default student-t(3,0,2.5) are too wide to translate

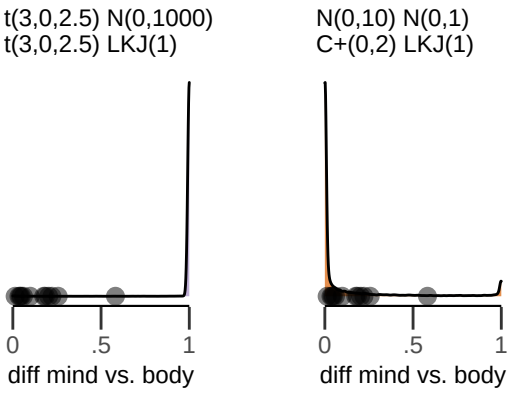
into reasonable predictions on the probability scale. Finally, a normal distribution with a standard deviation of 1 seems to make the best predictions for the experimental state effect, putting most mass on smaller differences, but still allowing for effects up to 75% (as observed in one previous study). Based on these prior predictions, we decided to use the normal(0,1) prior for the intercept and the effect, the half-normal(0,1) for the variation between countries, and the LKJ(2) for the correlation matrix.




N(0,5) N(0,1)
IG(3,0.5) LKJ(2)
N(0,1) N(0,1)
IG(3,0.5) LKJ(2)
N(0,5) N(0,0.5)
IG(3,0.5) LKJ(2)
N(0,1) N(0,0.5)
IG(3,0.5) LKJ(2)
N(0,5) N(0,5)
C+(0,2) LKJ(2)
N(0,1) N(0,5)
C+(0,2) LKJ(2)
N(0,5) N(0,1)
C+(0,2) LKJ(2)
N(0,1) N(0,1)
C+(0,2) LKJ(2)
N(0,5) N(0,0.5)
C+(0,2) LKJ(2)
N(0,1) N(0,0.5)
C+(0,2) LKJ(2)
N(0,5) N(0,5)
N+(0,1) LKJ(2)
N(0,1) N(0,5)
N+(0,1) LKJ(2)
N(0,5) N(0,1)
N+(0,1) LKJ(2)
N(0,1) N(0,1)
N+(0,1) LKJ(2)
N(0,5) N(0,0.5)
N+(0,1) LKJ(2)
N(0,1) N(0,0.5)
N+(0,1) LKJ(2)
0
.5
1
P('continues')


t(3,0,2.5) N(0,1000)
t(3,0,2.5) LKJ(1)
0
.5
1
P('continues')
N(0,10) N(0,1)
C+(0,2) LKJ(1)
0
.5
1
P('continues')
N(0,5) N(0,5)
Exp(1) LKJ(2)
0
.5
1
diff mind vs. body
N(0,1) N(0,5)
Exp(1) LKJ(2)
0
.5
1
diff mind vs. body
N(0,5) N(0,1)
Exp(1) LKJ(2)
0
.5
1
diff mind vs. body
N(0,1) N(0,1)
Exp(1) LKJ(2)
0
.5
1
diff mind vs. body
N(0,5) N(0,0.5)
Exp(1) LKJ(2)
0
.5
1
diff mind vs. body
N(0,1) N(0,0.5)
Exp(1) LKJ(2)
0
.5
1
diff mind vs. body
N(0,5) N(0,5)
IG(3,0.5) LKJ(2)
0
.5
1
diff mind vs. body
N(0,1) N(0,5)
IG(3,0.5) LKJ(2)
0
.5
1
diff mind vs. body


N(0,5) N(0,1)
IG(3,0.5) LKJ(2)
N(0,1) N(0,1)
IG(3,0.5) LKJ(2)
N(0,5) N(0,0.5)
IG(3,0.5) LKJ(2)
N(0,1) N(0,0.5)
IG(3,0.5) LKJ(2)
N(0,5) N(0,5)
C+(0,2) LKJ(2)
N(0,1) N(0,5)
C+(0,2) LKJ(2)
N(0,5) N(0,1)
C+(0,2) LKJ(2)
N(0,1) N(0,1)
C+(0,2) LKJ(2)
N(0,5) N(0,0.5)
C+(0,2) LKJ(2)
N(0,1) N(0,0.5)
C+(0,2) LKJ(2)
N(0,5) N(0,5)
N+(0,1) LKJ(2)
N(0,1) N(0,5)
N+(0,1) LKJ(2)
N(0,5) N(0,1)
N+(0,1) LKJ(2)
N(0,1) N(0,1)
N+(0,1) LKJ(2)
N(0,5) N(0,0.5)
N+(0,1) LKJ(2)
N(0,1) N(0,0.5)
N+(0,1) LKJ(2)
0
.5
1
diff mind vs. body



As becomes evident in Figure 13, predictions from both our preregistered prior settings and the brms default settings are completely unrealistic; both predict that all responses with be either complete cessation or continuity. The brms default priors are much too wide, resulting in predicting an unlikely difference of 100% between conditions. The preregistered priors, on the other hand, predict a modest effect, but due to the wide prior on the variation between countries, this results in a very strong prediction of observing no effect. However, note that in this case, because we have so much data, the data will always outweigh the priors, resulting in a reasonable posterior distribution, regardless of the exact prior specifications (see robustness checks).

MCMC Diagnostics

To investigate convergence of the MCMC chains, we extracted the $\hat{R}$ values for all model parameters. The smallest and largest $\hat{R}$ values were 1.00 for the correlation between the slope of the state effect and the state-by-context effect and 1.00 for the individual level religiosity effect, respectively. The traceplots for these smallest and largest $\hat{R}$ values are shown in Figure 14a and b.

The ratio of effective samples versus total samples $\hat{N}_{\text{eff}}/N$ was calculated per parameter to assess to what extent autocorrelation in the chains reduces the certainty of the posterior estimates (Geyer, 2011). Ideally, $\hat{N}_{\text{eff}}$ is as large as possible (Vehtari, Gelman, Simpson, Carpenter, & Bürkner, 2021b). The $\hat{N}_{\text{eff}}/N$ for each of the 315 estimated parameters is

displayed in Figure 14c. Note that $\hat{N}_{\text{eff}}$ can be larger than the the total number of iterations (in this case: $N = 20000$) when the samples are anti-correlated or antithetical (Carpenter, 2018). The smallest $\hat{N}_{\text{eff}} = 3016$ for the overall intercept. For many parameters, $\hat{N}_{\text{eff}}$ is at least half of the number of iterations, although for some parameters the ratio is rather low, indicating that there is some autocorrelation in the chains. Nevertheless, since brms uses the NUTS sampler (Hoffman & Gelman, 2014), even for complex models 'a few thousand' samples generally suffice for stable results (Bürkner, 2017b). We therefore concluded that the effective sample size is sufficient for valid interpretation of the estimates and inference.

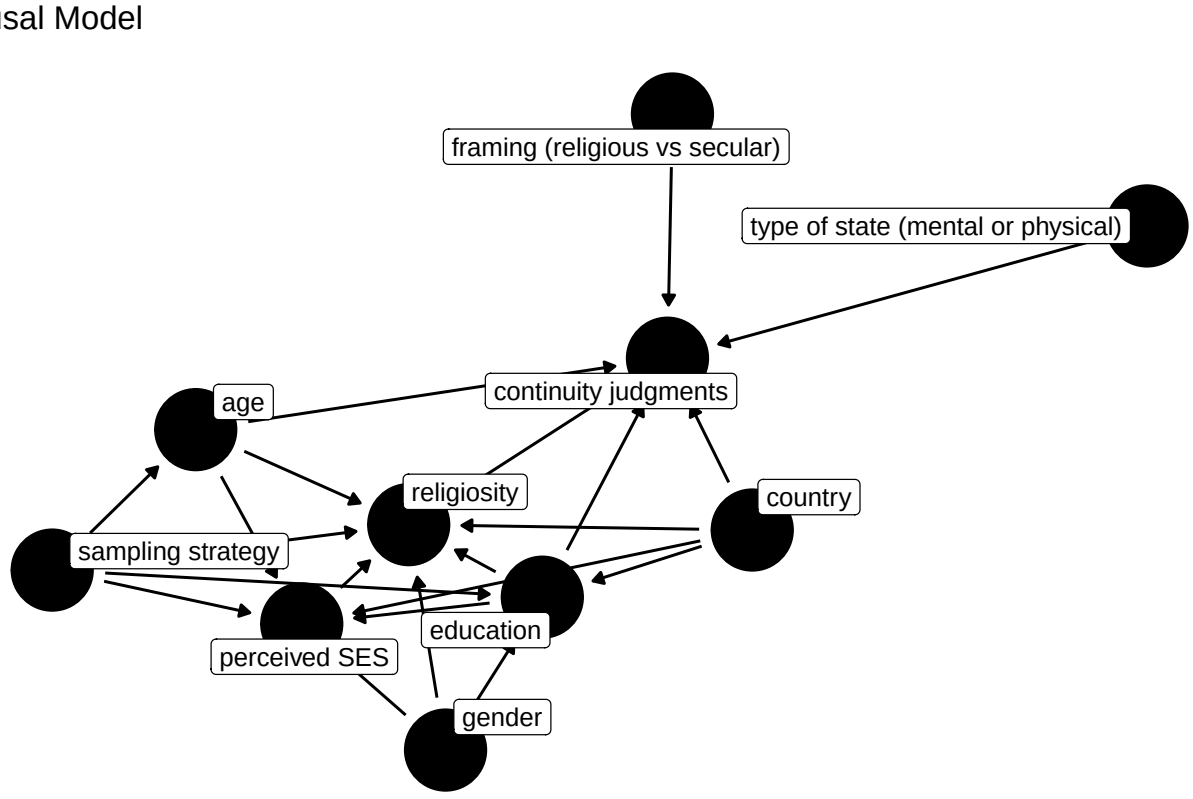
Causal assumptions

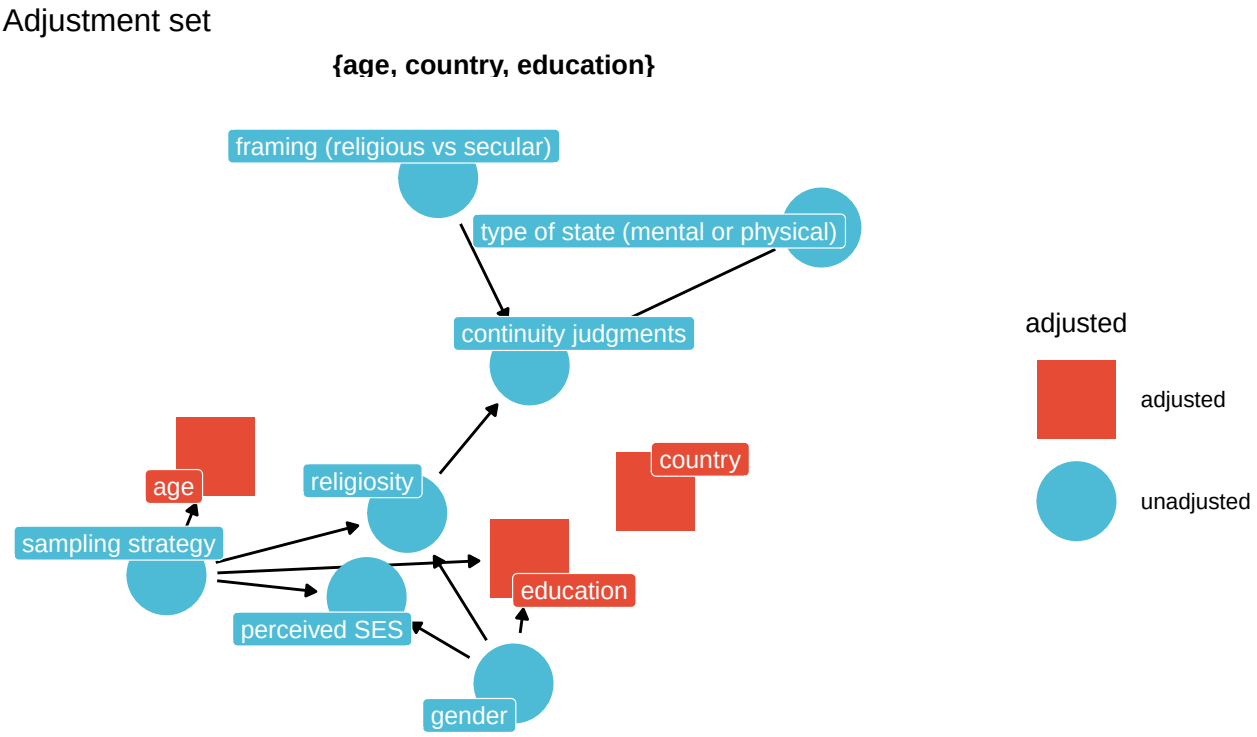
In order to systematically investigate which third variables should and should not be included in the statistical model, we used graphical causal models representing the relations between the variables in our data. As part of the data of interest is observational (e.g., religiosity, demographics), it is important to identify potential confounder variables, 'back-door paths', mediators and colliders that may affect causal inference (McElreath, 2016; Pearl, 2019; Rohrer, 2018). We identified the following structure based on theoretical assumptions about the measured variables:

- continuity judgments of states of a deceased individual are potentially affected by:
  - the type of state (mental vs. physical)
  - the framing of the narrative (religious vs. secular/medical)
  - country (culture)
  - education (knowledge of neuroscientific views on death/ mind-body relation)
  - religion
  - age (relevance of and personal experience with death of close friends/relatives)
- religion is affected by age, SES, education, gender, and country, sampling strategy
- SES is affected by country, education, age, and gender, sampling strategy
- education is affected by country, age, and gender, sampling strategy

- age is affected by sampling strategy
- **The exposure of interest is "religion"**
- **The outcome of interest is "continuity judgments"**

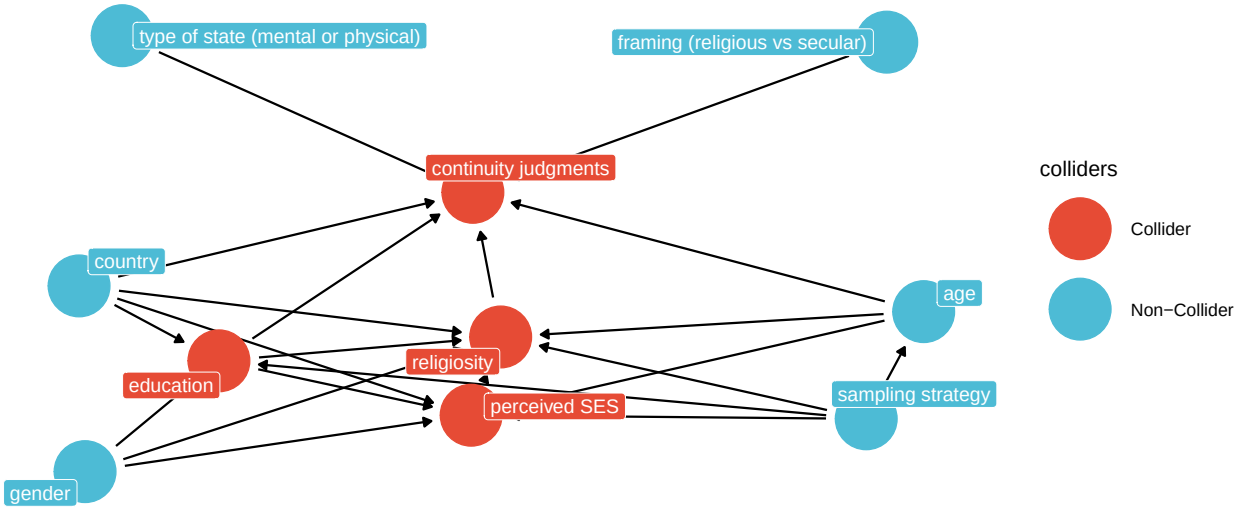
Using *directed acyclic graphs* (DAGs; Pearl (1995)) created in the R package `ggdag` (Barrett, 2021), this resulted in the following structure:



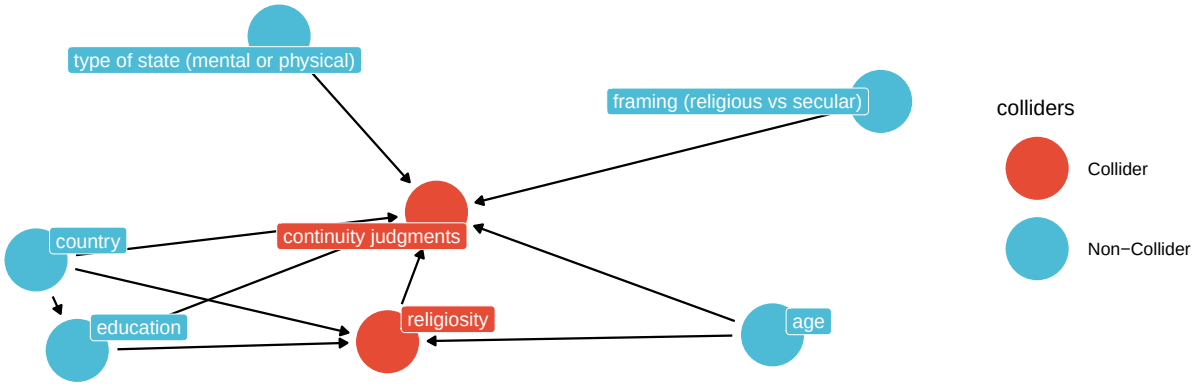


Thus, conditional on the assumed model, we should include *country*, *education*, and *age* as covariates or adjustment variables. So, rather than 'controlling for' all indicators that could affect either the predictor or outcome of interest, we only adjusted for the indicators that are needed for causal inference. As drawn in Figure **??**, in the large model, many covariates are identified as colliders; including those may introduce spurious associations and bias the relation of interest between religiosity and continuity judgments. In the adjusted model, none of the remaining covariates are colliders, making conditioning on country, age, and education valid inference choices.

a Lurking Colliders Large Model



b Lurking Colliders Adjusted Model



References

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, *63*(4), 596–612. doi:10.1037/0022-3514.63.4.596

Auguie, B. (2017). *gridExtra: Miscellaneous functions for "grid" graphics*. Retrieved from https://CRAN.R-project.org/package=gridExtra

Aust, F., & Barth, M. (2020). *papaja: Create APA manuscripts with R Markdown*. Retrieved from https://github.com/crsh/papaja

Barrett, M. (2021). Ggdag: Analyze and create elegant directed acyclic graphs.

Bates, D., & Maechler, M. (2021). *Matrix: Sparse and dense matrix classes and methods*. Retrieved from https://CRAN.R-project.org/package=Matrix

Betancourt, M., Vehtari, A., & Gelman, A. (2015). Prior choice recommendations.

Bürkner, P.-C. (2017a). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28. doi:10.18637/jss.v080.i01

Bürkner, P.-C. (2017b). Brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28. doi:10.18637/jss.v080.i01

Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, *10*(1), 395–411. doi:10.32614/RJ-2018-017

Bürkner, P.-C. (2021). Bayesian item response modeling in R with brms and Stan. *Journal of Statistical Software*, *100*(5), 1–54. doi:10.18637/jss.v100.i05

Carpenter, B. (2018). We were measuring the speed of Stan incorrectly—it's faster than we thought in some cases due to antithetical sampling. *Statistical Modeling, Causal Inference, and Social Science*.

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., … Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*. doi:10.18637/jss.v076.i01

Eddelbuettel, D., & Balamuta, J. J. (2018). Extending extitR with extitC++: A

Brief Introduction to extitRcpp. *The American Statistician*, *72*(1), 28–36. doi:10.1080/00031305.2017.1375990

Eddelbuettel, D., & François, R. (2011). Rcpp: Seamless R and C++ integration. *Journal of Statistical Software*, *40*(8), 1–18. doi:10.18637/jss.v040.i08

Forstmann, M., Burgmer, P., & Mussweiler, T. (2012). "The Mind Is Willing, but the Flesh Is Weak": The Effects of Mind-Body Dualism on Health Behavior. *Psychological Science*, *23*(10), 1239–1245. doi:10.1177/0956797612442392

Gabry, J., & Češnovar, R. (2021). *Cmdstanr: R interface to 'CmdStan'*.

Geyer, C. J. (2011). Introduction to Markov Chain Monte Carlo. In S. Brooks, A. Gelman, G. L. Jones, & X.-L. Meng (Eds.), *Handbook of Markov Chain Monte Carlo*. CRC Press. doi:10.1201/b10905-3

Gronau, Q. F., Singmann, H., & Wagenmakers, E.-J. (2020a). bridgesampling: An R package for estimating normalizing constants. *Journal of Statistical Software*, *92*(10), 1–29. doi:10.18637/jss.v092.i10

Gronau, Q. F., Singmann, H., & Wagenmakers, E.-J. (2020b). Bridgesampling: An R Package for Estimating Normalizing Constants. *Journal of Statistical Software*, *92*.

Haaf, J. M., & Rouder, J. N. (2017). Developing constraint in Bayesian mixed models. *Psychological Methods*, *22*(4), 779–798. doi:10.31234/osf.io/ktjnq

Harris, P. L., & Giménez, M. (2005). Children's Acceptance of Conflicting Testimony: The Case of Death. *Journal of Cognition and Culture*, *5*(1), 143–164. doi:10.1163/1568537054068606

Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, *15*, 1593–1623.

Hoogeveen, S., Haaf, J. M., Bulbulia, J. A., Ross, R. M., McKay, R., Altay, S., ... van Elk, M. (2022). The Einstein effect provides global evidence for scientific source credibility effects and the influence of religiosity. *Nature Human Behaviour*.

doi:10.1038/s41562-021-01273-8

Hox, J. J. C. M., van de Schoot, R., & Matthijsse, S. (2012). How few countries will do? Comparative survey analysis from a Bayesian perspective. *Survey Research Methods*, *6*, 87–93. doi:10.18148/srm/2012.v6i2.5033

Jeppson, H., Hofmann, H., & Cook, D. (2021). *Ggmosaic: Mosaic plots in the 'ggplot2' framework*. Retrieved from https://CRAN.R-project.org/package=ggmosaic

Kahle, D., & Stamey, J. (2017). *Invgamma: The inverse gamma distribution*. Retrieved from https://CRAN.R-project.org/package=invgamma

Kay, M. (2021). *ggdist: Visualizations of distributions and uncertainty*. doi:10.5281/zenodo.3879620

Kay, M. (2022). *tidybayes: Tidy data and geoms for Bayesian models*. doi:10.5281/zenodo.1308151

Lindeman, M., Svedholm-Hakkinen, A. M., & Lipsanen, J. (2015). Ontological confusions but not mentalizing abilities predict religious belief, paranormal belief, and belief in supernatural purpose. *Cognition*, *134*, 63–76. doi:10.1016/j.cognition.2014.09.008

Lun, V. M.-C., & Bond, M. H. (2013). Examining the relation of religion and spirituality to subjective well-being across national cultures. *Psychology of Religion and Spirituality*, *5*(4), 304–315. doi:10.1037/a0033641

Martin, A. D., Quinn, K. M., & Park, J. H. (2011). MCMCpack: Markov chain monte carlo in R. *Journal of Statistical Software*, *42*(9), 22. doi:10.18637/jss.v042.i09

McElreath, R. (2016). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. Boca Raton (FL): Chapman & Hall/CRC Press. doi:10.1201/9781315372495

Morey, R. D., & Rouder, J. N. (2021). *BayesFactor: Computation of bayes factors for common designs*. Retrieved from

https://CRAN.R-project.org/package=BayesFactor

Pearl, J. (1995). Causal diagrams for empirical research. *Biometrika*, *82*(4), 669–688. doi:10.1093/biomet/82.4.669

Pearl, J. (2019). The seven tools of causal inference, with reflections on machine learning. *Communications of the ACM*, *62*(3), 54—60. doi:10.1145/3241036

Pfadt, J. M., & van den Bergh, D. (2020). Bayesrel: Bayesian reliability estimation.

Plummer, M., Best, N., Cowles, K., & Vines, K. (2006). CODA: Convergence diagnosis and output analysis for MCMC. *R News*, *6*(1), 7–11. Retrieved from https://journal.r-project.org/archive/

R Core Team. (2021). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rohrer, J. M. (2018). Thinking clearly about correlations and causation: Graphical causal models for observational data. *Advances in Methods and Practices in Psychological Science*, *1*(1), 27–42. doi:10.1177/2515245917745629

Rouder, J. N., Haaf, J. M., Davis-Stober, C. P., & Hilgard, J. (2019). Beyond overall effects: A Bayesian approach to finding constraints in meta-analysis. *Psychological Methods*, *24*, 606–621. doi:10.1037/met0000216

Stan Development Team. (2020). StanHeaders: Headers for the R interface to Stan. Retrieved from https://mc-stan.org/

Stan Development Team. (2021). RStan: The R interface to Stan. Retrieved from https://mc-stan.org/

Stavrova, O. (2015). Religion, self-rated health, and mortality: Whether religiosity delays death depends on the cultural context. *Social Psychological and Personality Science*, *6*(8), 911–922. doi:10.1177/1948550615593149

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., & Bürkner, P.-C. (2021a). Rank-normalization, folding, and localization: An improved rhat for assessing

convergence of MCMC (with discussion). *Bayesian Analysis*.

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., & Bürkner, P.-C. (2021b). Rank-Normalization, Folding, and Localization: An Improved R^ for Assessing Convergence of MCMC. *Bayesian Analysis*, 1–28. doi:10.1214/20-BA1221

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with s* (Fourth.). New York: Springer. Retrieved from https://www.stats.ox.ac.uk/pub/MASS4/

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. Retrieved from https://ggplot2.tidyverse.org

Wickham, H. (2021). *Tidyr: Tidy messy data*. Retrieved from https://CRAN.R-project.org/package=tidyr

Wickham, H., François, R., Henry, L., & Müller, K. (2021). *Dplyr: A grammar of data manipulation*. Retrieved from https://CRAN.R-project.org/package=dplyr

Wickham, H., & Seidel, D. (2020). *Scales: Scale functions for visualization*. Retrieved from https://CRAN.R-project.org/package=scales

Wilke, C. O. (2020). *Cowplot: Streamlined plot theme and plot annotations for 'ggplot2'*. Retrieved from https://CRAN.R-project.org/package=cowplot

World Values Survey. (2010). Wave 6 Official Aggregate v. 20150418. worldvaluessurvey.org; Asep/JDS Madrid.

Xiao, N. (2018). *Ggsci: Scientific journal and sci-fi themed color palettes for 'ggplot2'*. Retrieved from https://CRAN.R-project.org/package=ggsci

Zhu, H. (2021). *kableExtra: Construct complex table with 'kable' and pipe syntax*. Retrieved from https://CRAN.R-project.org/package=kableExtra

Table 1
Descriptive Statistics per Country

Country	N	Age (SD)	Women	Religiosity	Sample	Compensation
Australia	463	48.3 (16.0)	48.4%	0.52	online panel	money
Belgium	320	34.6 (13.1)	55.6%	0.24	mixed	raffle
Brazil	402	28.8 (10.4)	73.1%	0.51	mixed	none; credits
Canada	351	33.2 (10.5)	52.4%	0.28	online panel	money
Chile	308	30.8 (9.9)	59.1%	0.33	mixed	raffle
China	390	32.1 (8.4)	55.9%	0.32	online panel	money
Croatia	309	28.0 (6.9)	78.3%	0.41	mixed	raffle
Denmark	415	27.9 (10.3)	71.3%	0.26	mixed	raffle
France	405	40.6 (12.8)	64.2%	0.29	online panel	money
Germany	1,287	27.5 (9.0)	62.2%	0.32	mixed	raffle
India	394	30.4 (6.5)	36.3%	0.73	online panel	money
Ireland	434	42.6 (15.0)	51.8%	0.48	online panel	money
Israel	501	27.9 (10.1)	73.5%	0.37	students	credits
Italy	342	27.2 (8.2)	50.9%	0.26	mixed	none; money
Japan	424	40.6 (10.0)	43.9%	0.29	online panel	money
Lithuania	291	24.1 (7.0)	83.2%	0.35	students	none
Morocco	329	32.1 (11.8)	16.1%	0.70	online panel	money
Netherlands	482	57.6 (14.7)	25.3%	0.28	online panel	money
Romania	539	24.4 (7.4)	85.2%	0.55	mixed	raffle
Singapore	308	22.2 (3.4)	62.0%	0.45	students	credits
Spain	337	41.9 (13.9)	31.2%	0.21	online panel	money
Turkey	362	39.2 (11.1)	24.6%	0.33	online panel	money
UK	400	36.2 (12.7)	65.8%	0.23	online panel	money
US	402	35.8 (14.4)	51.0%	0.45	mixed	none; money
Total	10,195	33.8 (13.8)	55.9%	0.38	-	-

Note. Religiosity refers tot he self-reported level of individual religiosity based on 9 items, transformed on a 0-1 scale. Sample indicates the sample composition based on the method of recruitment per site.

Table 2
Religious Denomination per Country

	Religious group						
Country	Christian	Muslim	Hindu	Buddhist	Jewish	Other	None
Australia	44.3%	5.4%	0.2%	0.9%	0.4%	1.9%	46.9%
Belgium	28.4%	2.5%	0.0%	0.6%	0.3%	0.6%	67.5%
Brazil	30.1%	0.0%	0.0%	1.0%	0.2%	14.4%	54.2%
Canada	26.5%	1.1%	0.9%	1.1%	2.0%	1.4%	67.0%
Chile	25.6%	0.0%	0.6%	1.6%	3.2%	2.3%	66.6%
China	3.6%	0.0%	0.0%	10.5%	0.0%	1.0%	84.9%
Croatia	54.4%	0.3%	0.0%	0.6%	0.3%	0.6%	43.7%
Denmark	35.7%	2.2%	0.0%	0.0%	0.0%	0.0%	62.2%
France	38.8%	6.2%	0.0%	0.2%	0.0%	1.2%	53.6%
Germany	54.4%	3.3%	0.1%	0.1%	0.3%	1.2%	40.7%
India	13.2%	3.6%	60.4%	0.3%	0.3%	0.8%	21.6%
Ireland	54.4%	1.6%	0.2%	0.0%	0.2%	0.9%	42.6%
Israel	2.2%	3.2%	0.0%	0.0%	11.6%	2.0%	81.0%
Italy	17.5%	0.0%	0.0%	0.9%	0.0%	0.0%	81.6%
Japan	0.9%	0.2%	0.0%	15.3%	0.0%	1.2%	82.3%
Lithuania	39.2%	0.0%	0.3%	0.0%	0.0%	0.7%	59.8%
Morocco	0.3%	78.1%	0.0%	0.0%	0.0%	1.5%	20.1%
Netherlands	27.0%	0.0%	0.0%	0.0%	0.6%	3.1%	69.3%
Romania	77.2%	0.2%	0.0%	0.2%	0.2%	2.2%	20.0%
Singapore	20.5%	4.9%	3.9%	20.5%	0.0%	5.2%	45.1%
Spain	39.8%	0.0%	0.0%	0.0%	0.0%	1.2%	59.1%
Turkey	0.0%	42.5%	0.0%	0.0%	0.3%	2.5%	54.7%
UK	22.2%	0.5%	0.8%	0.5%	0.8%	1.0%	74.2%
US	44.0%	1.2%	0.5%	0.2%	3.2%	3.0%	47.8%
Total	32.0%	5.7%	2.6%	2.0%	1.0%	2.0%	54.6%

Note. Percentage of people indicating to be member of the respective religious groups. Note that the response options were particularized per country. Here we show the 5 most prevalent groups.

Table 3
Bayes factor model comparison and parameter estimates for the key effects

	Bayes factors				Parameter estimates	
Effect	$\mathcal{M}_0$	$\mathcal{M}_1$	$\mathcal{M}_+$	$\mathcal{M}_u$	μ	σ
State Effect	0.00	0.00	**1.00**	0.09	1.71 [1.55, 1.86]	0.35 [0.25, 0.50]
Religiosity Effect	0.00	0.00	**1.00**	0.09	0.84 [0.71, 0.96]	0.28 [0.21, 0.39]
Framing Effect	0.00	0.00	**1.00**	0.09	0.52 [0.41, 0.61]	0.22 [0.15, 0.32]
State-by-Religiosity Effect	0.00	0.00	0.84	**1.00**	0.24 [0.14, 0.33]	0.18 [0.11, 0.28]
State-by-Framing Effect	**1.00**	0.02	0.00	0.11	-0.09 [-0.19, 0.00]	0.08 [0.00, 0.22]

Note. The preferred model for each effect is assigned value 1.00 and displayed in bold. The remaining values are the Bayes factors for the respective model relative to this preferred model. Subscripts reflect constraints on the critical parameter; $_0$ indicates no effect, $_1$ indicates a common (positive) effect, $_+$ indicates a varying positive effect, and $_u$ indicates an unconstrained effect. Parameter estimates (median and 95% credible interval) are taken from the unconstrained model for $\mathcal{H}_5$.

Table 4
Bayes factor of different models for robustness checks

Robustness set	μ [95% CI]	BF_{10}	BF_{+1}	Preferred
State Effect				
Main analysis	1.71 [1.55, 1.86]	∞	10^{26}	$\mathcal{M}_+$
Hearing as mental	1.16 [1.04, 1.27]	∞	10^{11}	$\mathcal{M}_+$
Manipulation check failure	1.75 [1.59, 1.88]	∞	10^{20}	$\mathcal{M}_+$
Education and age as covariates	1.73 [1.57, 1.88]	∞	10^{26}	$\mathcal{M}_+$
Excluding Lithuania[a]	1.70 [1.54, 1.85]	∞	10^{26}	$\mathcal{M}_+$
Cauchy prior on sd[a]	1.71 [1.54, 1.86]	∞	10^{25}	$\mathcal{M}_+$
Religiosity Effect				
Main analysis	0.84 [0.71, 0.96]	∞	10^{87}	$\mathcal{M}_+$
Hearing as mental	0.88 [0.75, 1.01]	∞	10^{116}	$\mathcal{M}_+$
Manipulation check failure	0.85 [0.72, 0.96]	∞	10^{75}	$\mathcal{M}_+$
Education and age as covariates	0.86 [0.73, 0.97]	∞	10^{80}	$\mathcal{M}_+$
Excluding Lithuania[a]	0.83 [0.70, 0.94]	∞	10^{85}	$\mathcal{M}_+$
Cauchy prior on sd[a]	0.84 [0.71, 0.96]	∞	10^{87}	$\mathcal{M}_+$
Context Effect				
Main analysis	0.52 [0.41, 0.61]	10^{135}	10^{11}	$\mathcal{M}_+$
Hearing as mental	0.56 [0.45, 0.66]	10^{174}	10^{16}	$\mathcal{M}_+$
Manipulation check failure	0.52 [0.41, 0.63]	10^{127}	10^{14}	$\mathcal{M}_+$
Education and age as covariates	0.51 [0.41, 0.61]	10^{133}	10^{10}	$\mathcal{M}_+$
Excluding Lithuania[a]	0.52 [0.41, 0.62]	10^{132}	10^{11}	$\mathcal{M}_+$
Cauchy prior on sd[a]	0.51 [0.41, 0.62]	10^{135}	10^{10}	$\mathcal{M}_+$
State-by-Religiosity Effect				
Main analysis	0.24 [0.14, 0.33]	10^{16}	3131	$\mathcal{M}_u$
Hearing as mental	0.14 [0.05, 0.22]	10^{5}	7.69	$\mathcal{M}_u$
Manipulation check failure	0.24 [0.14, 0.33]	10^{15}	180	$\mathcal{M}_+$
Education and age as covariates	0.24 [0.14, 0.34]	10^{17}	6342	$\mathcal{M}_u$
Excluding Lithuania[a]	0.23 [0.13, 0.33]	10^{15}	2576	$\mathcal{M}_u$
Cauchy prior on sd[a]	0.24 [0.14, 0.33]	10^{16}	1158	$\mathcal{M}_u$
State-by-Context Effect				
Main analysis	-0.09 [-0.19, 0.00]	0.02	0.04	$\mathcal{M}_0$
Hearing as mental	-0.11 [-0.20, -0.02]	0.02	0.03	$\mathcal{M}_0$
Manipulation check failure	-0.07 [-0.17, 0.02]	0.03	0.03	$\mathcal{M}_0$
Education and age as covariates	-0.09 [-0.18, 0.01]	0.03	0.02	$\mathcal{M}_0$
Excluding Lithuania[a]	-0.10 [-0.19, -0.01]	0.02	0.03	$\mathcal{M}_0$
Cauchy prior on sd[a]	-0.09 [-0.18, 0.01]	0.03	0.01	$\mathcal{M}_0$

Note:
Across all five sets of robustness checks, the results is qualitatively equal to those of the main analyses; the data indicate (a) strong state, religiosity, and context effects that vary between countries but are consistently positive (mind > body; religious > non-religious; theistic context > secular context), (b) a varying state-by-religiosity interaction effect (though sometimes the unconstrained model is preferred), and (c) no state-by-context effect. Subscripts reflect parameter constraints; $_0$ indicates the null model, $_+$ indicates a varying positive effect, and $_1$ indicates a common effect.

Table 5
Bayes factor model comparison and parameter estimates for the key effects for atheists extinctivists only

	Bayes factors				Parameter estimates	
Effect	$\mathcal{M}_0$	$\mathcal{M}_1$	$\mathcal{M}_+$	$\mathcal{M}_u$	μ	σ
State Effect	0.00	0.06	**1.00**	0.09	1.63 [1.30, 1.93]	0.44 [0.12, 0.85]
Framing Effect	0.00	0.03	**1.00**	0.45	0.78 [0.44, 1.10]	0.56 [0.23, 0.96]
State-by-Framing Effect	0.69	0.09	0.02	**1.00**	-0.39 [-0.85, 0.07]	0.28 [0.01, 0.89]

Note. The preferred model for each effect is assigned value 1.00 and displayed in bold. The remaining values are the Bayes factors for the respective model relative to this preferred model. Subscripts reflect constraints on the critical parameter; $_0$ indicates no effect, $_1$ indicates a common (positive) effect, $_+$ indicates a varying positive effect, and $_u$ indicates an unconstrained effect. Parameter estimates (median and 95% credible interval) are taken from the unconstrained model for $\mathcal{H}_5$.

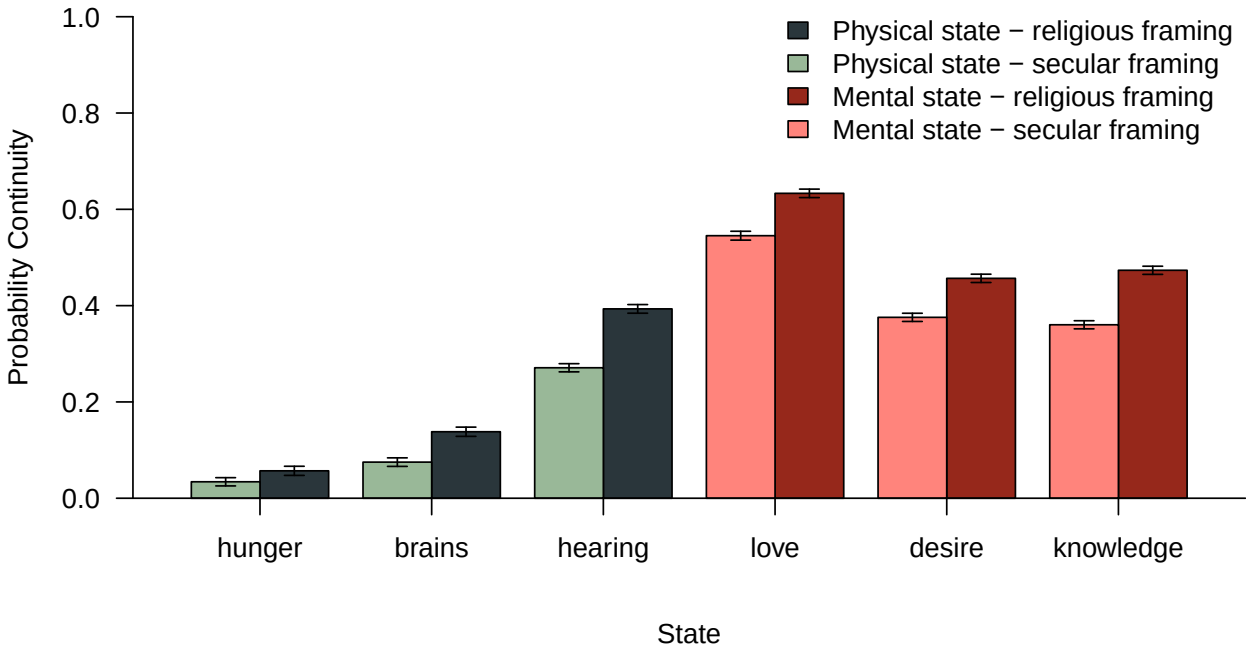


Figure 1. Descriptive pattern per state. Probability of continuity judgment per item, displayed per framing condition. The states were measured within-subjects and the framing was manipulated between-subjects.

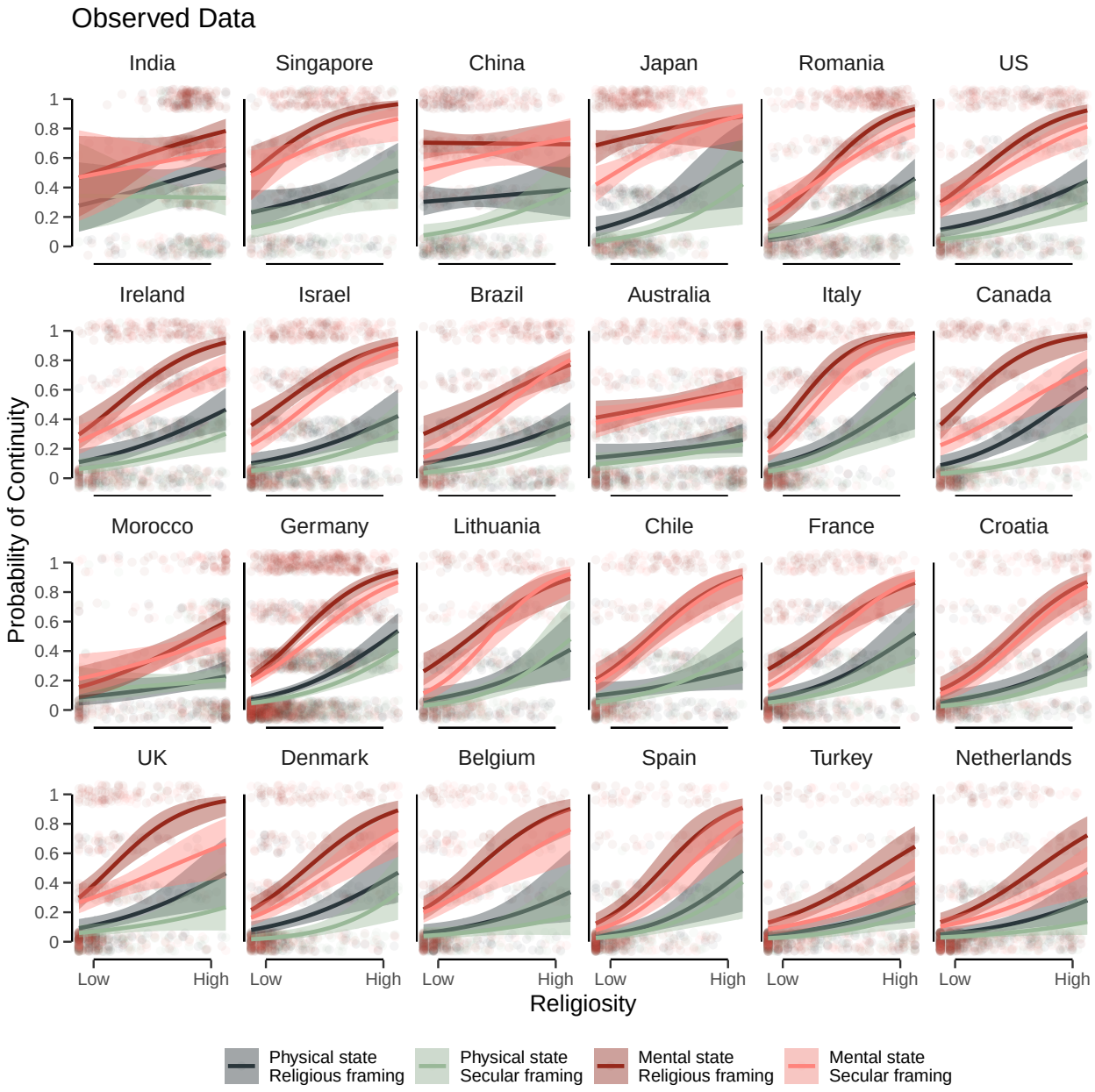


Figure 2. Descriptive pattern of results per country. Countries are ordered by the overall probability of making a continuity judgment (from left to right, top to bottom). Red lines denote probabilities for mental states and blue lines denote probabilities for the physical states. Solid lines denote probabilities in the religious context and dashed lines denote probabilities in the secular context. Data points are jittered to enhance visibility. Probabilities are averaged over the three items per category and religiosity was split into 6 levels.

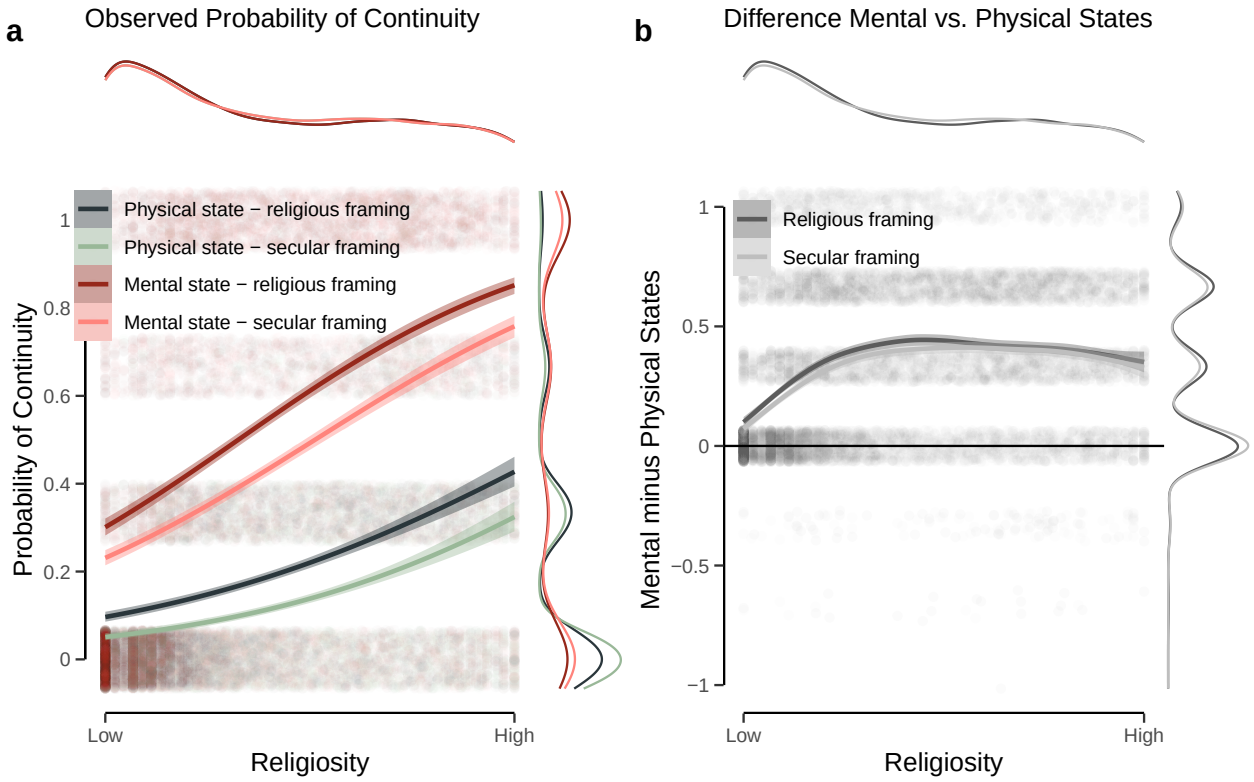


Figure 3. Descriptive pattern of results. Panel **a** displays the probability of making a continuity judgment per state category (physical vs. mental), framing (secular vs. religious) and individual level of religiosity. Panel **b** shows the difference in probability for mental vs. physical processes (contrast effect).

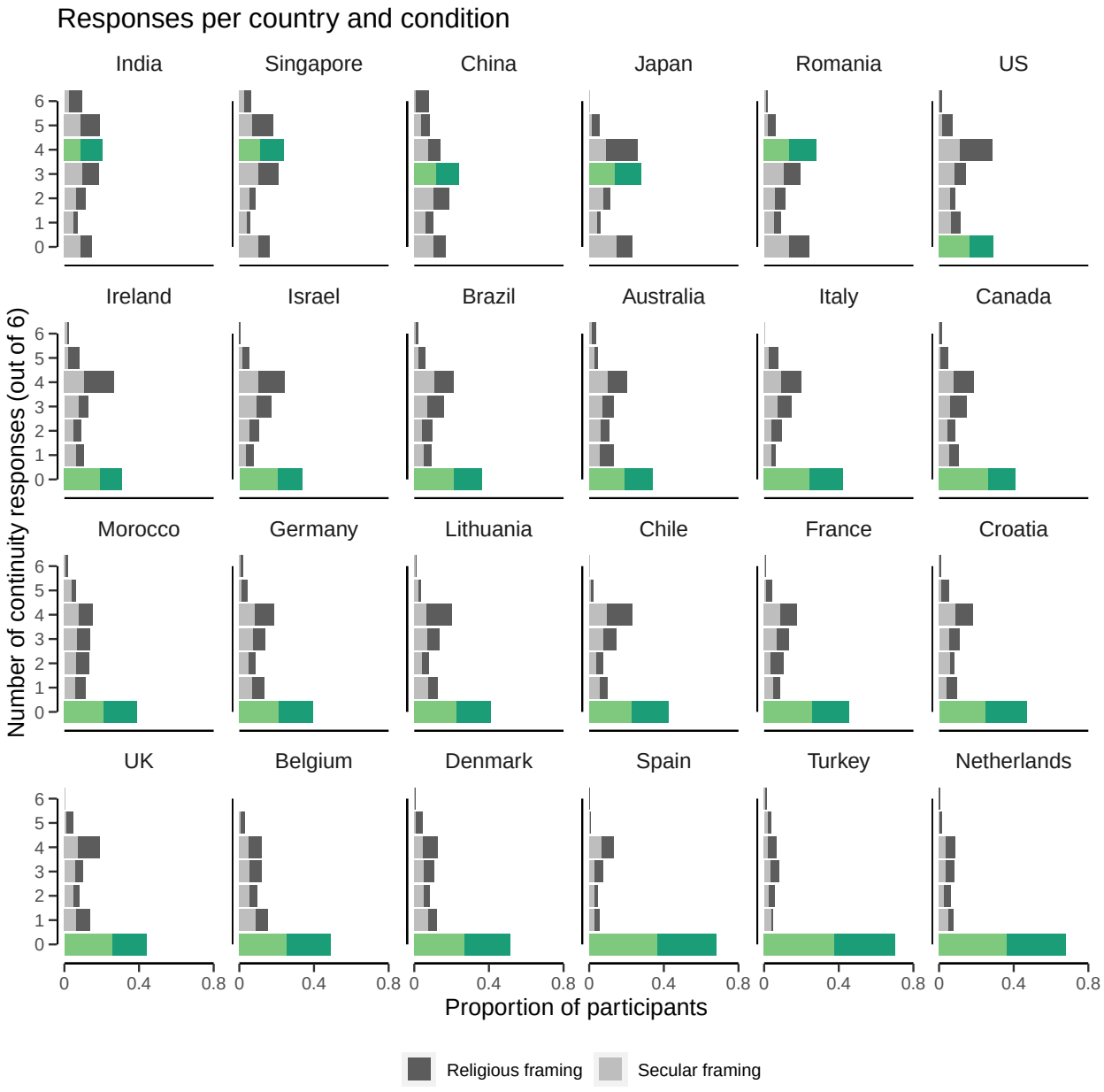


Figure 4. Proportion of participants and number of continuity responses per country. Countries are ordered by the overall probability of making a continuity judgment (from left to right, top to bottom). Dark grey bars reflect responses in the religious context and light grey bars reflect responses in the secular context. The modal number of continuity responses per country is indicated in green. Continuity responses were out of 6 states.

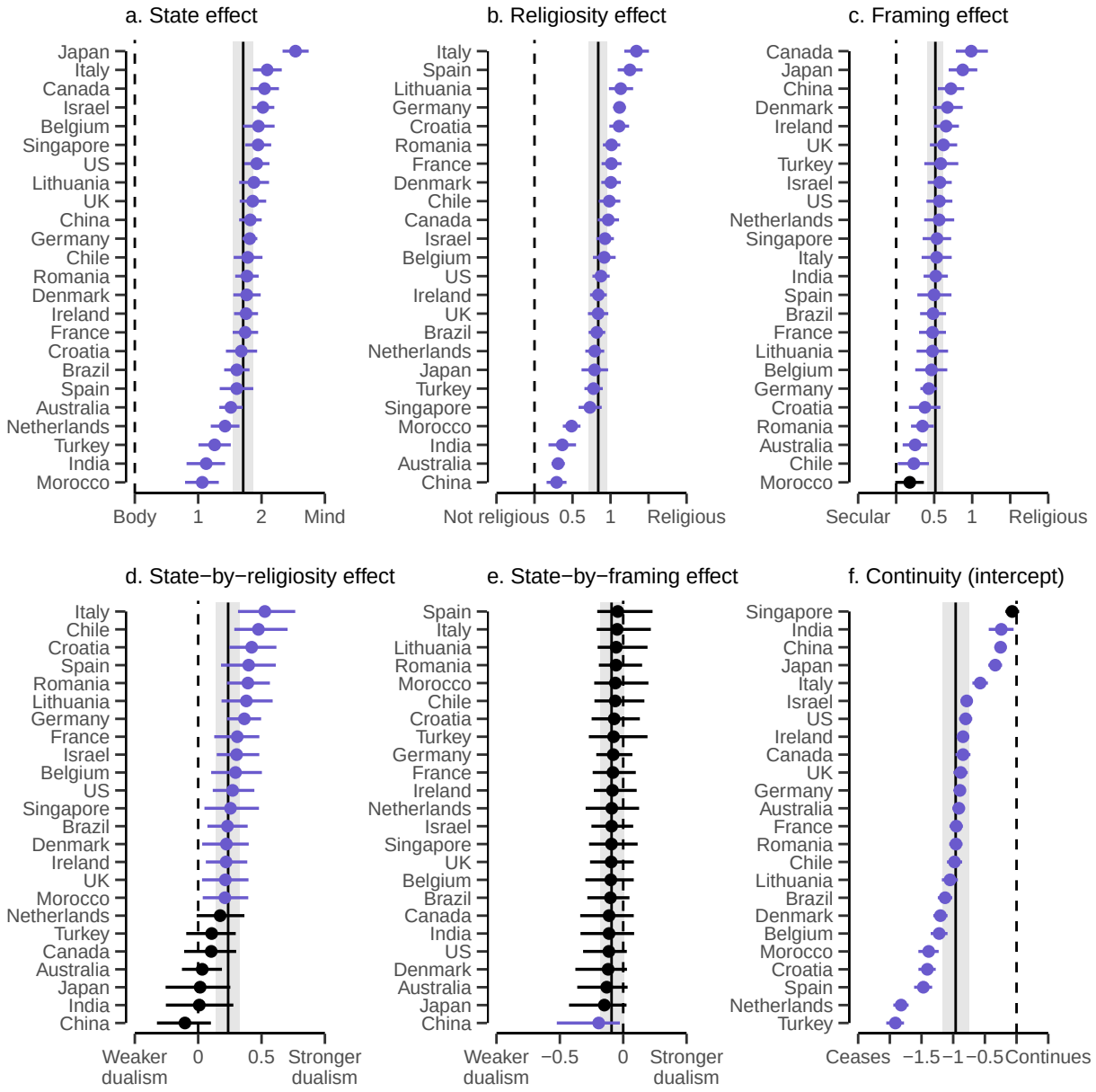


Figure 5. Estimated country-level effects (posterior medians) in increasing order. **a.** state contrast effects. **b.** religiosity effects. **c.** framing effects. **d.** state-by-religiosity interaction effects. **e.** state-by-framing interaction effects. **f.** intercepts. Each dot represents a country. Estimates with credible intervals colored in red exclude zero and estimates with credible intervals colored in grey include zero. The errorbars give the 95% credible interval for each country. The vertical lines denote the posterior median of the overall mean of the respective effect with the 95% credible interval in the shaded bands. The dashed lines indicates zero.

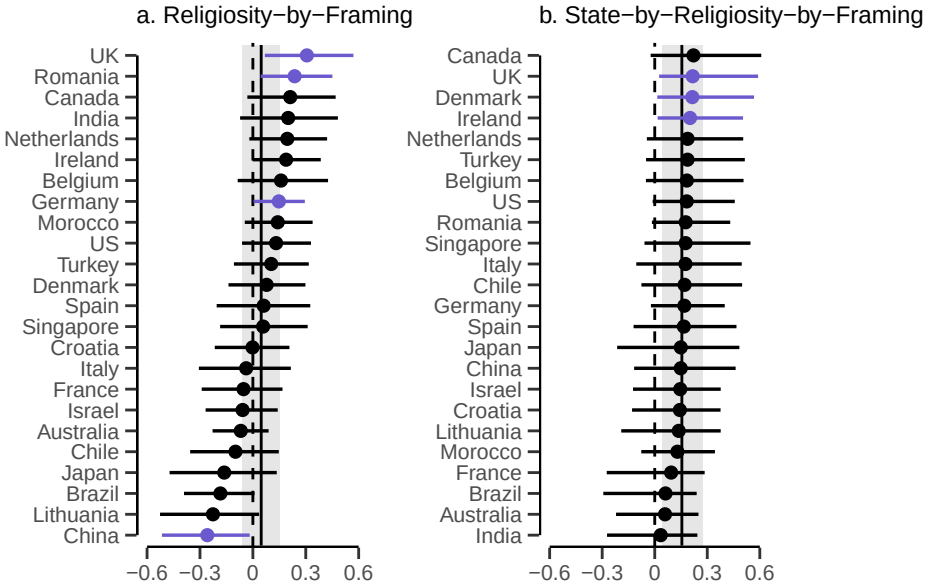


Figure 6. Estimated country-level effects (posterior medians) in increasing order. a. Religion-by-context interaction effects, where positive values indicate more continuity judgments for religious individuals in the religious framing condition. b. State-by-religiosity-by-context threeway interaction effects, where positive values indicate stronger mind-body dualism for religious individuals in the religious framing context. Each dot represents a country. Estimates with credible intervals colored in purple exclude zero and estimates with credible intervals colored in black include zero. The errorbars give the 95% credible interval for each country. The vertical lines denote the posterior median of the overall mean of the respective effect with the 95% credible interval in the shaded bands. The dashed lines indicates zero.

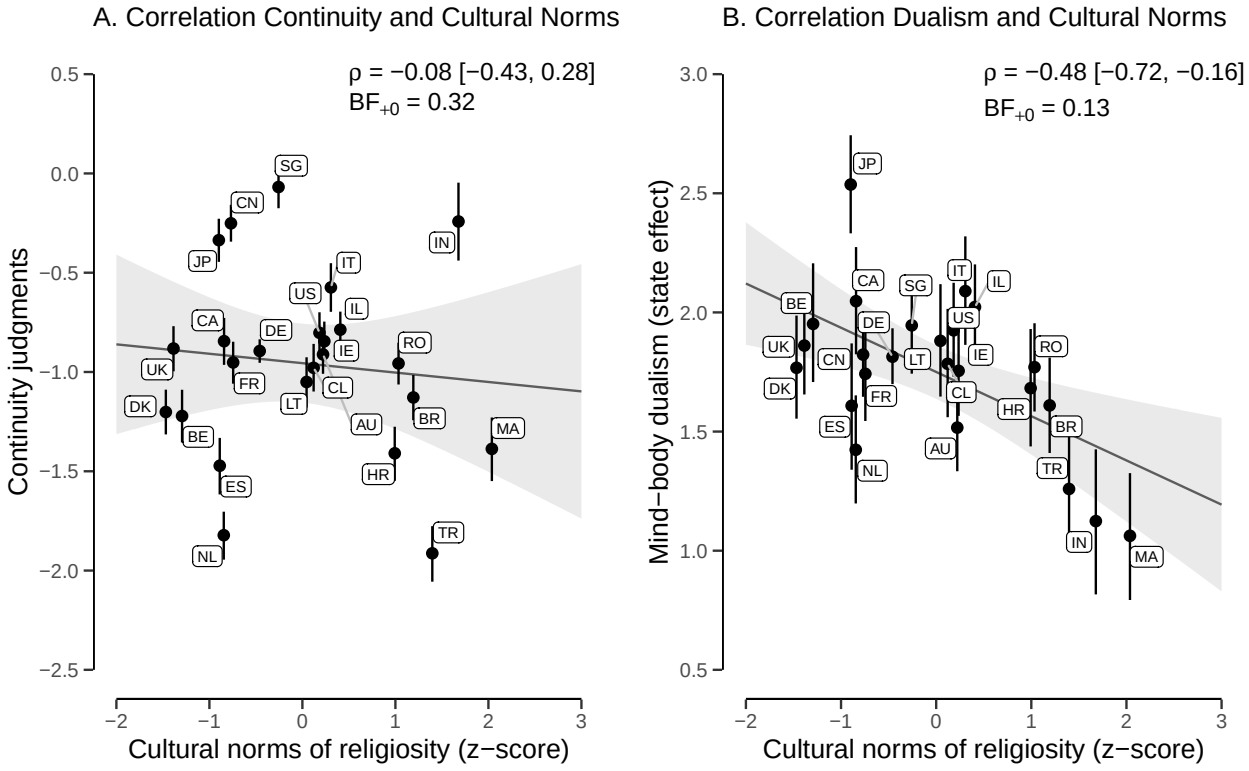


Figure 7. Correlation between country-level cultural norms of religion and continuity judgments (panel A.) and mind-body dualism (i.e., state effects; panel B.).

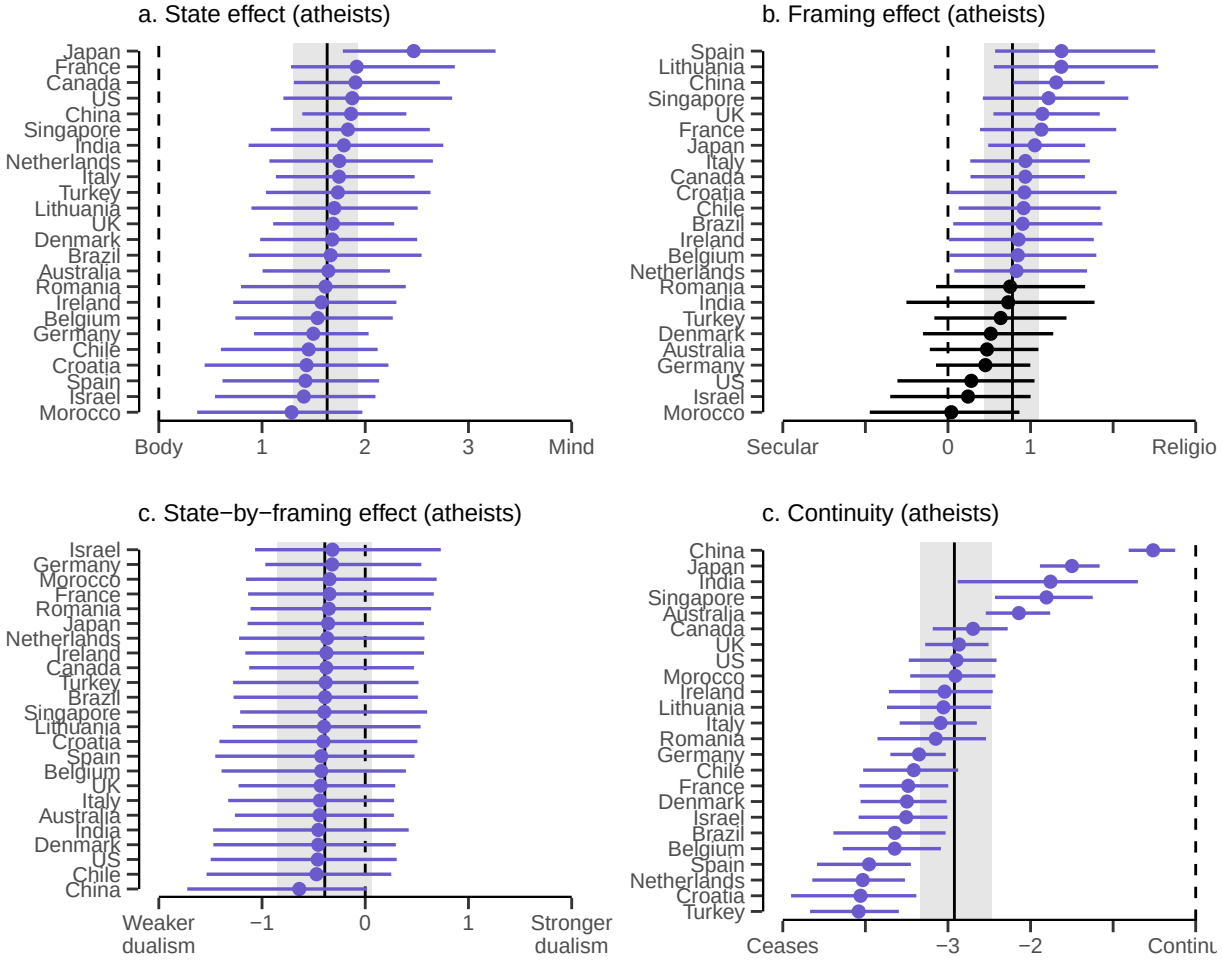


Figure 8. Estimated country-level effects (posterior medians) in increasing order. a. state contrast effects. b. framing effects. c. state-by-framing effects. d. intercepts. Each dot represents a country. Estimates with credible intervals colored in red exclude zero and estimates with credible intervals colored in grey include zero. The errorbars give the 95% credible interval for each country. The vertical lines denote the posterior median of the overall mean of the respective effect with the 95% credible interval in the shaded bands. The dashed lines indicates zero.

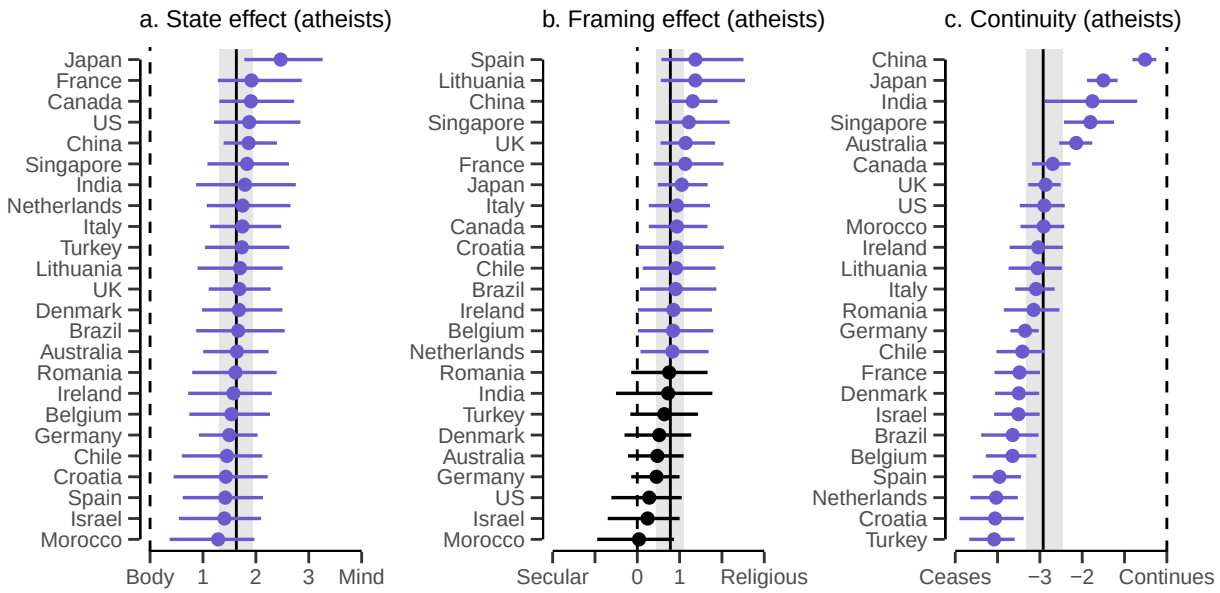


Figure 9. Estimated country-level effects (posterior medians) in increasing order. a. state contrast effects. b. framing effects. c. state-by-framing effects. d. intercepts. Each dot represents a country. Estimates with credible intervals colored in red exclude zero and estimates with credible intervals colored in grey include zero. The errorbars give the 95% credible interval for each country. The vertical lines denote the posterior median of the overall mean of the respective effect with the 95% credible interval in the shaded bands. The dashed lines indicates zero.

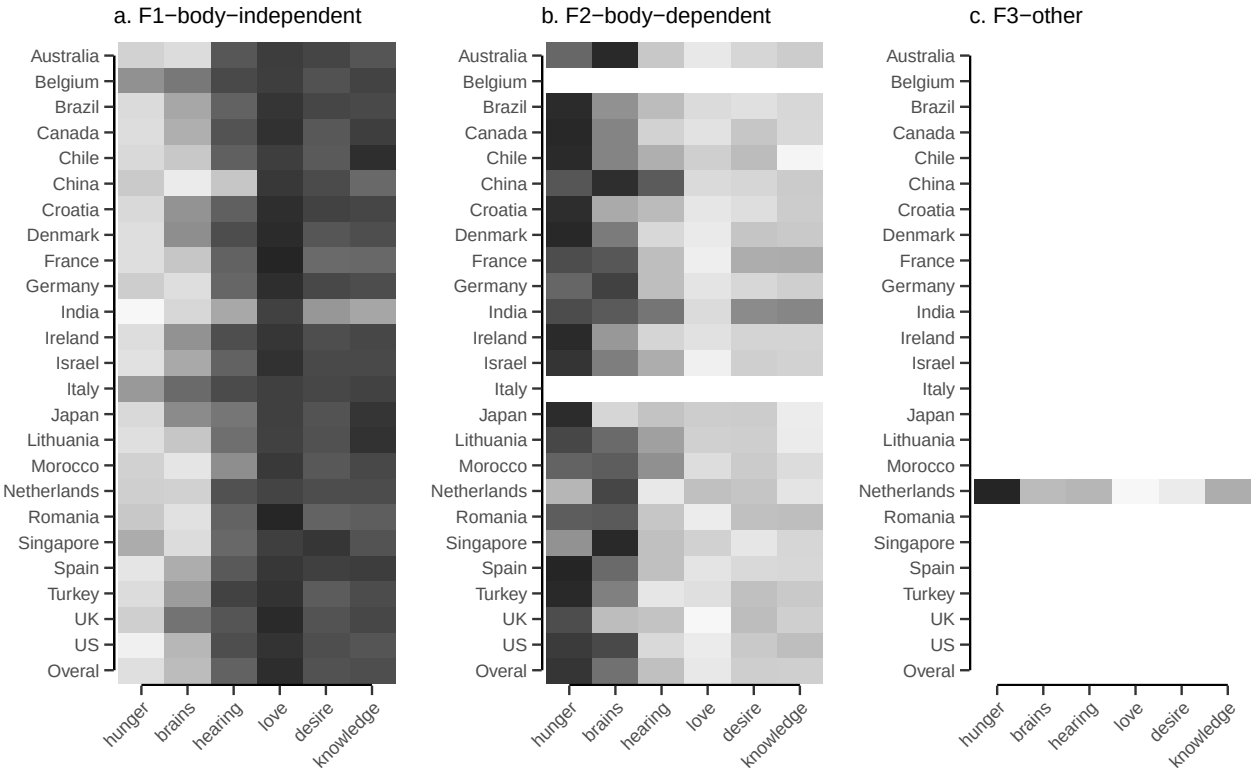


Figure 10. Factor loadings from EFAs per country. a. represents body-independent or mind-like factors, b. represents body-dependent or body-like factors and c. represents other factors. Factors were extracted using tetrachoric correlations.

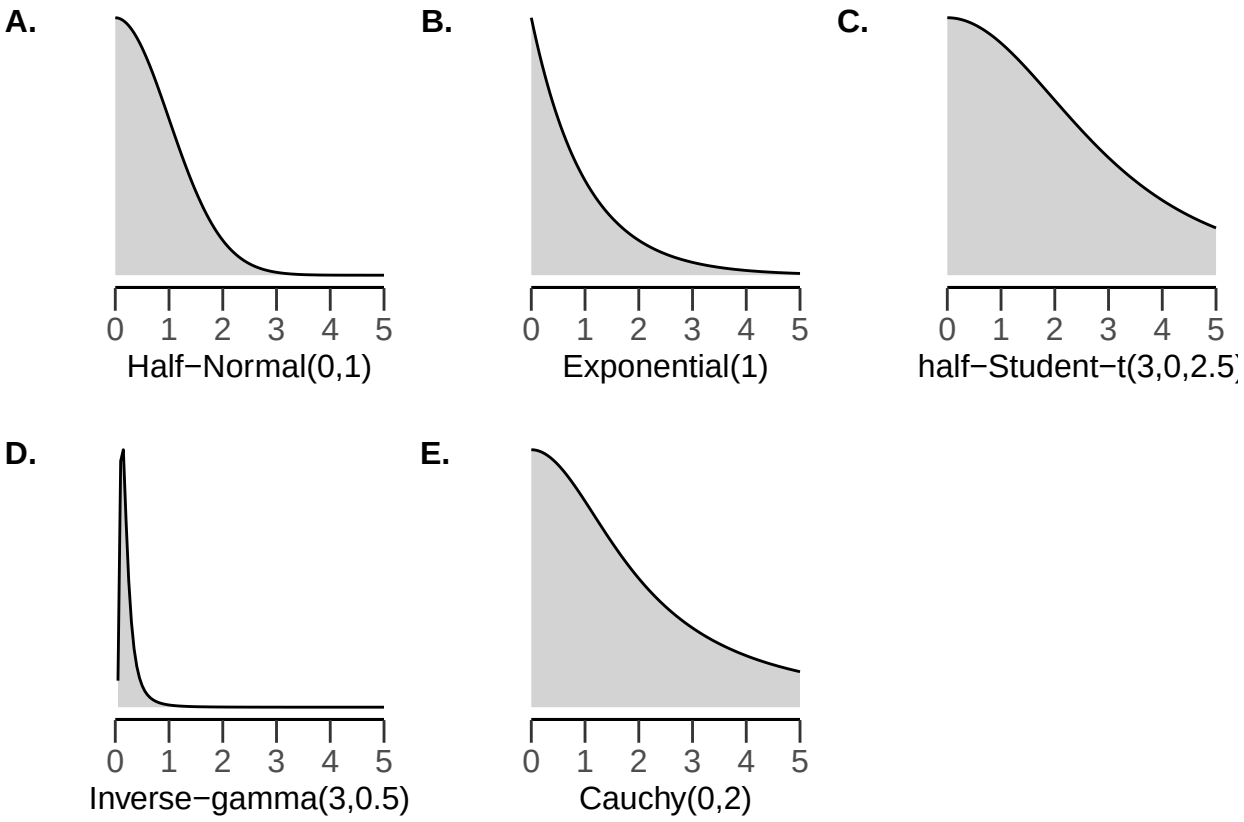


Figure 11. Different prior settings for the between-country variation in the effects of interest.

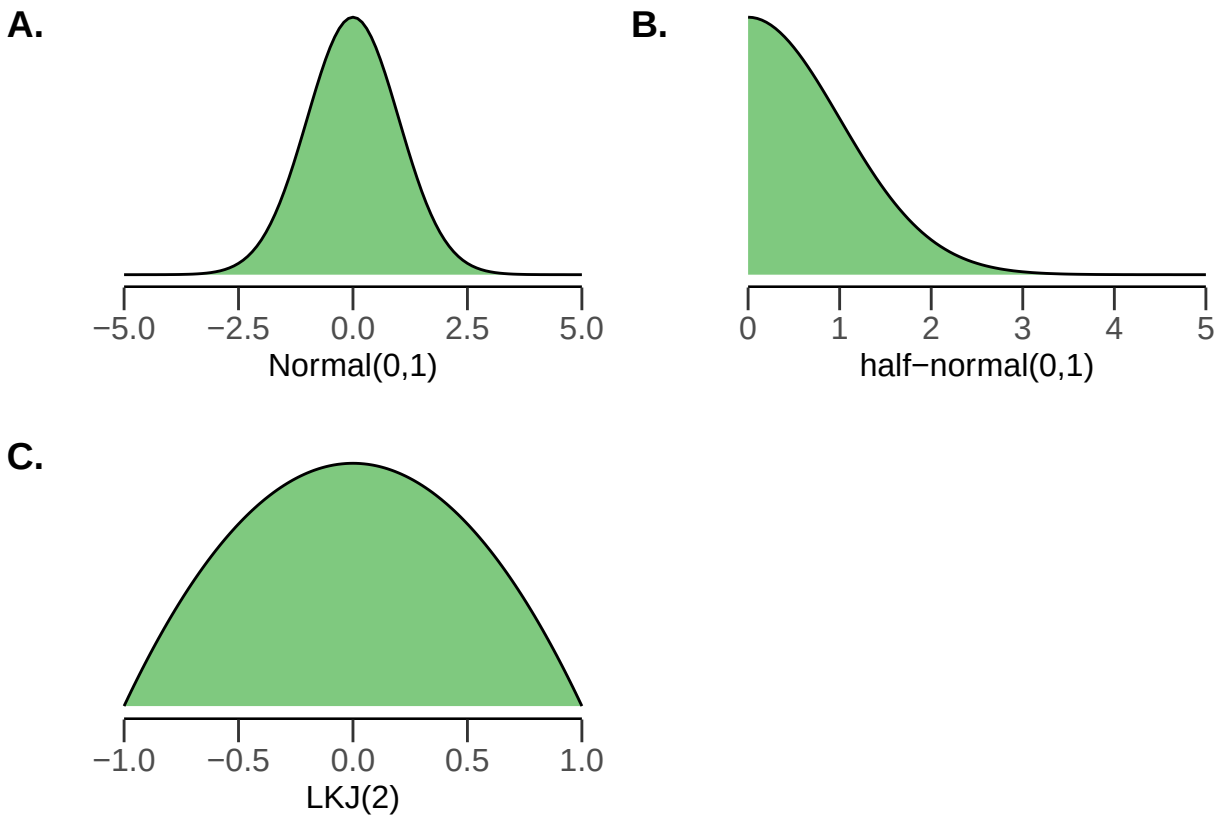


Figure 12. Chosen prior settings for the main analysis. A. shows the prior on the intercept and the effect, B. shows the prior on the variability between countries, and C. shows the prior on the correlation matrix.

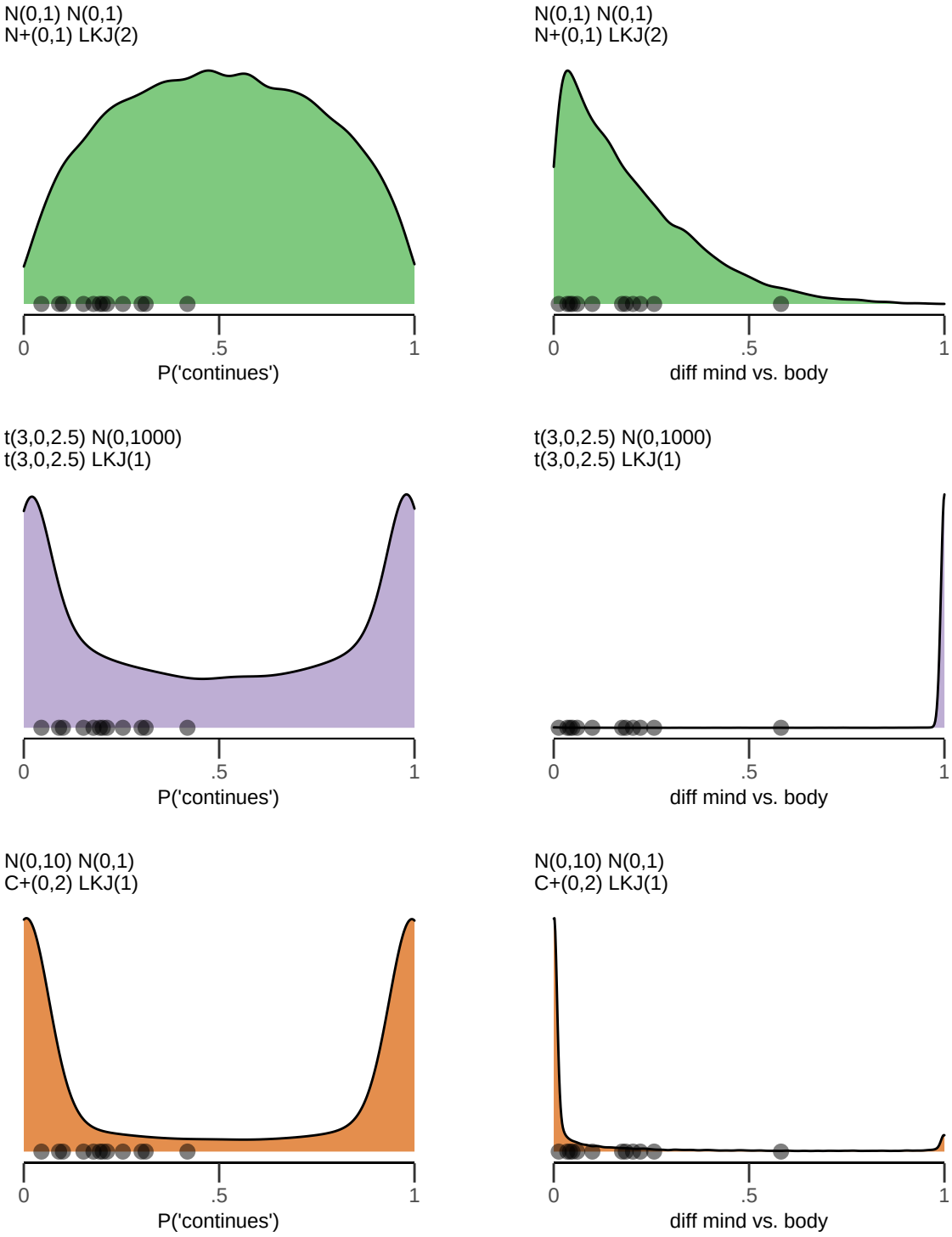


Figure 13. Prior predictions for the chosen prior settings (in green), the brms default settings (in purple), and the preregistered settings (in orange).

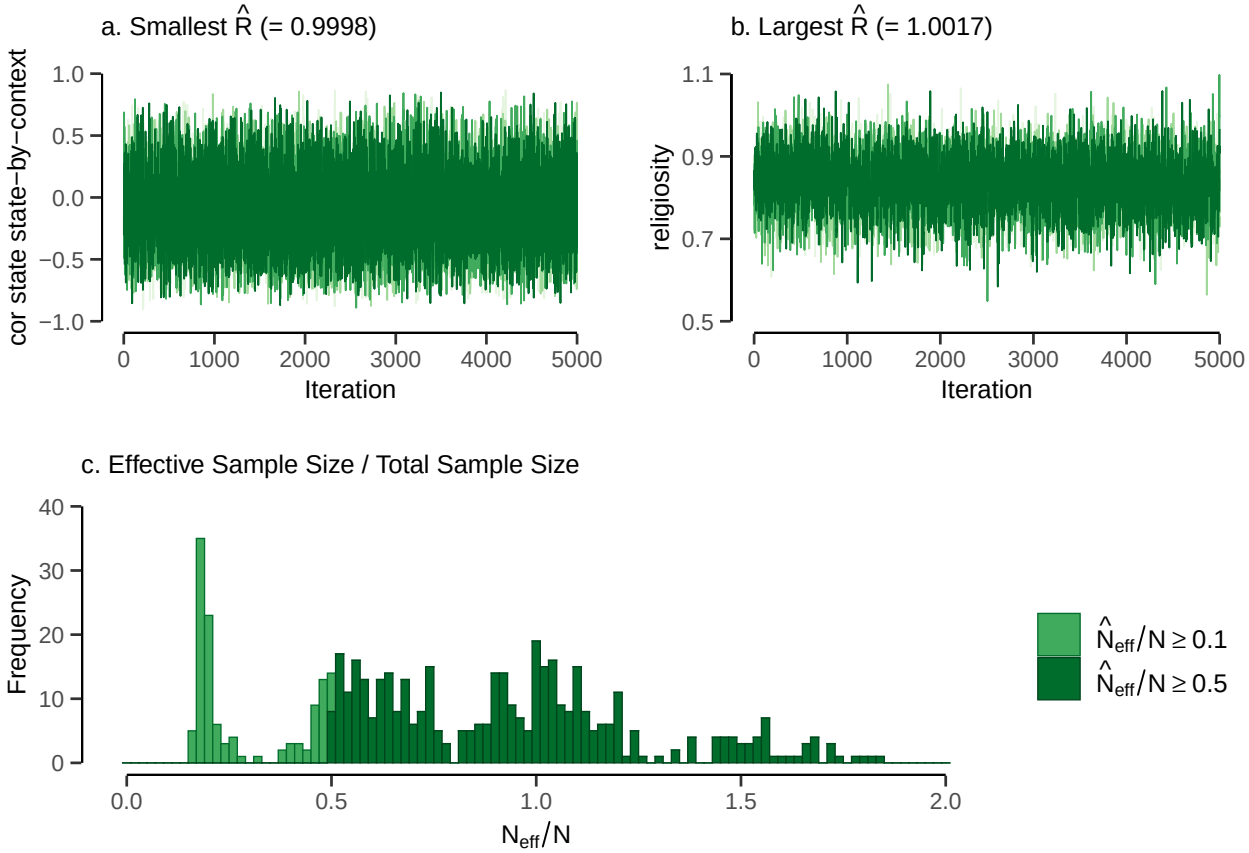


Figure 14. MCMC diagnostics. a. Chains for parameters with the smallest (correlation between the slope for the state effect and the state-by-context interaction effect) and b. largest (individual level religiosity effect) rhat values. c. Ratio of the number of effective samples versus the total samples for each parameter in the full model.