Gods, Games and the Socioecological Landscape (GGSL) Project: Core Causal Model for Site-Specific Papers (Objective 1)

Daniel Major-Smith[1,2], Augusto Dalla Ragione[3], Sevgi Demiroglu[4], Byambabaatar Ichinkhorloo[5], Martin Kocsis[3], Attila Mátéffy[6], Anders Norge[7,8], Connor Wood[9], and Benjamin Grant Purzycki[1]

[1]Department of the Study of Religion, Aarhus University, Aarhus, Denmark
[2]Population Health Sciences: Bristol Medical School, University of Bristol, Bristol, UK
[3]Department of Human Behavior, Ecology and Culture, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany
[4]Department of Anthropology, University of Connecticut, Connecticut, USA
[5]Institute for Mongolia Studies, National University of Mongolia, Ulaan Baatar, Mongolia
[6]Department of Mongolian and Tibetan Studies, University of Bonn, Bonn, Germany
[7]Department of Cultural Anthropology and Ethnology, Uppsala University, Uppsala, Sweden
[8]Department of Anthropology, Aarhus University, Aarhus, Denmark
[9]Center for Mind and Culture, Boston University, Boston, Massachusetts, USA

February 4, 2026

1 Introduction

Objective 1 of the Gods, Games and the Socioecological Landscape (GGSL) project aims to understand the putative causal relationship between religious beliefs - chiefly belief in punitive and omniscient deities - and the expansion of human cooperation with geographically distant co-religionists. Briefly, this project is theoretically motivated by the long-held idea that religion and religious beliefs may perform an adaptive social function by promoting cooperation, particularly among distant group members beyond kinship and reciprocal partnerships (see e.g., Johnson and Krüger, 2004; Norenzayan et al., 2016; Purzycki and Sosis, 2022). A key proposed cognitive mechanism to encourage this cooperative behaviour is belief in punitive and omniscient supernatural agents (the 'supernatural punishment' hypothesis; Johnson, 2005). Although various mechanisms have been proposed (Schloss and Murray, 2011), broadly, this theory states that by increasing the perceived benefits of cooperative behaviour and/or increasing the perceived costs of punishment for defection/selfish behaviour, the threat of supernatural sanctions may expand the scope of cooperation beyond our everyday interaction spheres.

The current project builds upon the previous Evolution of Religion and Morality (ERM) project, which reported a positive relationship between belief in supernatural punishment and omniscient gods and experimental measures of cooperation with distant co-religionists across 15 small-scale societies (Purzycki et al., 2016; Lang et al., 2019; Purzycki et al., 2022). Despite these previous findings, there are numerous open questions which the current project aims to examine (for further discussion, see Purzycki et al., 2024). First, as the previous studies predominantly focused on 'moralistic' religious traditions (i.e., Christianity, Hinduism, etc.), the extent to which these findings may apply to local, indigenous and less-explicitly-moralistic (henceforth 'local') gods is unknown. We address this in the present GGSL project by focusing on the relationship between religious beliefs regarding *local* deities and cooperation with distant co-religionists of the same local religious traditions. Second, previous studies did not take an explicitly causal framework to try and answer the question of whether beliefs in punitive and omniscient supernatural agents *cause* increased cooperation towards distant co-religionists. While causal inference from observational data is fraught with difficulties and relies on many, often unverifiable, assumptions (e.g., no confounding, selection or measurement biases), being clear about the causal nature of the research question, the estimand(s) of interest, and the assumptions required for a causal interpretation is vital for communicating and interpreting results (McElreath, 2020; Pearl et al., 2016; Hernán and Robins, 2020; Rohrer, 2018).

The aim of of Objective 1 of the GGSL project is therefore to examine whether beliefs regarding supernatural punishment and omniscience of local or indigenous deities (i.e., spirits from non-world religions) may promote cooperation among distant co-religionists of the same local religious tradition, using explicitly causal methods. We pursue this goal using site-specific (the present document) and omnibus (pre-registration forthcoming) analyses. Below, we: i) briefly introduce the field sites; ii) describe our core causal model for our site-specific analyses; iii) briefly introduce the data collection methods; and iv) describe our core statistical model based on our core causal model.

2 Field Sites

Data for the GGSL project (described in more detail below) were/are being collected in seven field sites (see Figure 1). Briefly, these are:

- Altai Uriankhai, Western Mongolia, a predominantly herding group who combine Buddhist and spirit-master/shamanist religious practices (researcher: Attila Mátéffy). Moralistic God (MG): Buddha. Local God (LG): Spirit-masters.
- Buryat, Eastern Mongolia, a predominantly herding group who combine Buddhist and spirit-master/shamanist religious practices (researcher: Byamba Ichinkhorloo). MG: Buddha. LG: Spirit-masters.
- BaYaka, Congo, a hunter-gatherer group with indigenous deities (researcher: Martin Kocsis). MG: Komba (Creator diety). LG: Forest-spirits.
- Coastal Benin, two primarily fishing, trading, and salt producing villages with both traditionally Vodun (voodoo) religion and more recently-introduced Christianity (researcher: Augusto Della Ragione). MG: Christian God. LG: Vodun spirits.
- Kichwa, Ecuador, a predominantly horticultural group with both Christian and local spirit religious beliefs and practices (researcher: Connor Wood). MG: Christian God. LG: Pachamama (a local nature spirit).
- Sihanaka, Madagascar, a predominantly agricultural group with local Sihanakan religious beliefs and deities (researcher: Anders Norge). MG: Ndrianampanjaka (a local sovereign deity concerned with peoples' behaviour). LG: Zazavavindrano (a local water-based spirit).
- Toraja, Sulawesi (Indonesia), an agriculturalist, moderately market-integrated, group who hold indigenous *aluk* and Christian (and, to a lesser extent, Islamic) beliefs (researcher: Sevgi Demiroglu). MG: Christian God. LG: Local ancestor spirits.

3 Core Causal Model

The main goal of this document is to detail the core causal model which acts as a starting point for the tailored causal models in each of the site-specific reports of this project. Note that this document is not a standard pre-registration (e.g., as is common for many psychological studies), as data has already been collected or data collection is in progress. Note, however, that we specified our methods prior to data collection in the grant proposal which can be found here: `https://github.com/bgpurzycki/GGSL-Project`). Instead, the present aim is to present the core causal model, and corresponding statistical models, *prior to analyses*. These will then inform our site-specific reports.

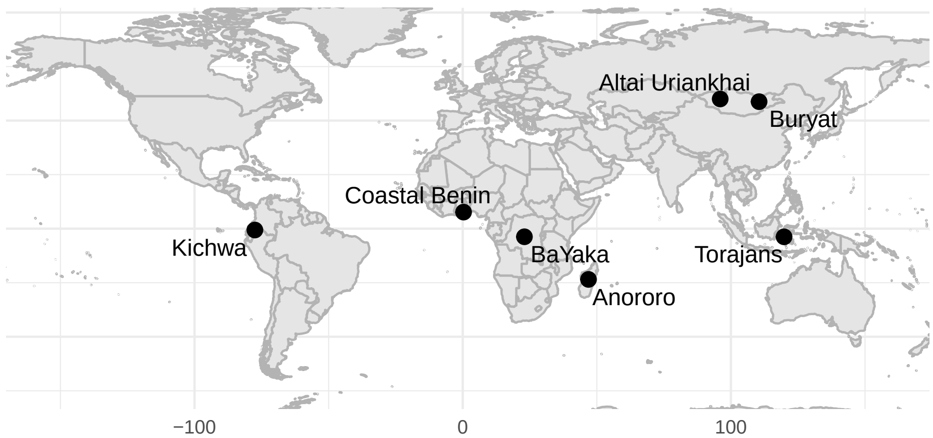

Figure 1: Map of GGSL Objective 1 field sites.

We present this causal model here to both act as a foundation for any subsequent site-specific reports to build upon and add contextually-relevant and ethnographically-informed detail. For instance, based on knowledge of their specific field site, researchers may wish to include additional variables as plausible confounders; equally, researchers may know from ethnographic experience that a certain variable is almost certainly not a confounder, and is plausibly a potential mediator. In these cases, we will update our causal model accordingly in each site-specific report and note any modifications. The aim of this core causal model is to have an initial model to work with, and act as a centralised pre-registration document for each of the site-specific papers of this project (note that a cross-cultural 'omnibus' model for a paper exploring the relationship between religious beliefs and cooperation with distant co-religionists across all GGSL and ERM societies, and whether this relationship differs by 'moralistic' vs 'local' religious traditions will be developed separately).

As discussed briefly above, there are a number of benefits to taking a causal inference approach to observational research (Hernán and Robins, 2020; Pearl et al., 2016; McElreath, 2020; Rohrer, 2018), including:

- *Having a clear research question:* Here, our research question is explicitly causal. Theory predicts that punitive and omniscient supernatural beliefs should *cause* cooperation, so this is what we are trying to estimate (as best as possible, given our assumptions)

- *Having a clear estimand* (i.e., the target quantity of interest we are trying to estimate; Lundberg et al., 2021): Our estimand from this causal model is "The joint average causal effect of supernatural punishment and omniscience beliefs regarding a local deity on experimental measures of cooperation with an unknown distant co-religionist of the same local religious tradition" (this will be unpacked in more detail below).

- *Clarity regarding the assumptions required to estimate an unbiased estimand:* These assumptions will be encoded in our Directed Acyclic Graph (DAG; see below), but essentially means that we assume: i) no confounding bias (i.e., all back-door paths between the exposure [religious beliefs] and the outcome [cooperation] are blocked, and no mediators or colliders are included in the adjustment set); ii) no selection bias (i.e., our sample is broadly representative of the target population, or that any non-representativeness is independent of our exposure and outcome); iii) no measurement bias (i.e., our variables are measured without error); and iv) no reverse causality (i.e., our outcome does not cause our exposure).

- *Assessing impacts of assumption violations and facilitating critique and discussion:* By making our research question, estimand(s) and assumptions for a causal interpretation explicit - and, importantly, discussing where assumptions may be violated and how this may impact any subsequent causal interpretations - readers are in a better position to evaluate and interpret the results of our study (e.g., is this assumption valid? What about X as another potential confounder?).

Our core causal assumptions for this project are presented in the DAG below (Figure 2). We'll now walk through this DAG and its assumptions.

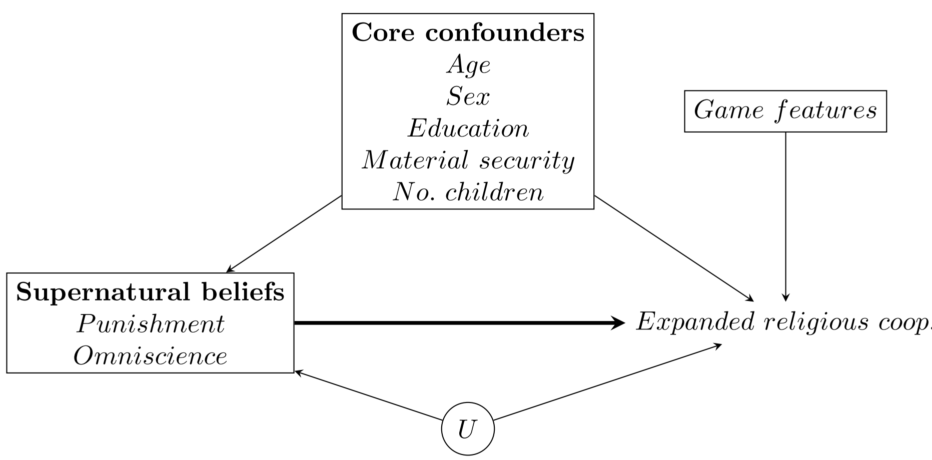

Figure 2: Directed Acyclic Graph (DAG) displaying our core causal assumptions regarding the relationship between supernatural beliefs (our exposure) and religious cooperation (our outcome).

First, our exposure is the joint effect of beliefs of local gods' punishment and omniscience. As data are cross-sectional and the causal relations between punishment and omniscience beliefs are likely to be complex and bidirectional, we focus on the *joint* effect of these variables. That is, we side-step the issue of any dynamism between punishment beliefs and omniscience beliefs or whether they are co-caused by a latent 'religiosity' factor. A downside of treating these variables jointly is that we are unable to estimate the *independent* effects of punishment and omniscience; to do so would require even more unverifiable - and likely implausible - assumptions than we are willing to make. For instance, to estimate the causal effect of just supernatural punishment on cooperation, we would need to know whether omniscience causes punishment - in which case omniscience would plausibly be a confounder, and hence need including as a covariate in regression models - or whether punishment causes omniscience - in which case omniscience would plausibly be a mediator, and hence not need including as a covariate - or whether the relationship is more complex (e.g., reciprocal causation, causation by a joint latent 'religiosity' factor, or a heady combination of all of the above); given this, we choose to make fewer assumptions and model these supernatural belief effects jointly.

Second, there is a bold arrow going from our exposure to our outcome ('Expanded religious cooperation'); this is our estimand of interest. That is, we want to estimate the joint causal effect of belief in a punitive and knowledgeable local god on experimental measures of cooperation with distant co-religionists of the same local religious tradition. Importantly, from this DAG we are assuming no reverse causality; i.e., that cooperation towards geographically distant co-religionists does not cause supernatural religious beliefs. While cooperation in these games specifically cannot cause religious beliefs, it is possible that prior cooperation outside this experimental context may cause both supernatural beliefs and cooperation in these games, and hence may be an unmeasured confounder; while perhaps possible, based on theory and previous experience we believe that this causal arrow (i.e., cooperation causing supernatural beliefs) is less likely - or at least much weaker - than the causal directionality from supernatural beliefs to religious cooperation we are assuming here.

Third, there are five variables - age, sex, education, material security and number of children - which we designate as core confounders of this relationship. That is, these variables plausibly cause both the exposure and the outcome, hence the arrows from the 'core confounders' node to both the exposure and outcome. Given this, these variables need to be adjusted for/conditioned upon (e.g., included as covariates in multivariable regression models) in order to block any open back-door biasing paths between the exposure and outcome. While these are assumptions, based on previous research (e.g., Baimel et al., 2022; Vardy et al., 2022; Purzycki and Bendixen, 2024), logical considerations (e.g., supernatural beliefs cannot cause age or sex) and temporal precedence (e.g., education levels are fixed in adulthood, hence less likely to cause our exposures) we believe they are plausible. Even if they are not confounders - e.g., if sex only causes supernatural beliefs and age only causes cooperation, say - their inclusion as covariates is unlikely to result in any bias. We are also assuming that these variables are *causes of* the exposure rather than being *caused by* the exposure; that is, we assume they are confounders, rather than mediators, of the exposure-outcome relationship. This assumption is reasonable for age, sex and education, although there is room for doubt with material security and number of children, as these may be caused by supernatural religious beliefs, and not just a cause of them (although both directions are possible). We will discuss these issues with the field researchers prior to any analyses to make sure any assumptions are plausible given the local ethnographic context. If, for instance, we deem it plausible that number of children might be a mediator rather than a confounder (or both), we can explore different causal assumptions to see how this impacts our conclusions (for more details on this type of approach to sensitivity analyses to different causal assumptions, see Major-Smith, 2023).

Fourth, there is a node 'U' on the DAG, representing unmeasured confounding. That is, if there are any additional confounders we have not measured or included in our causal model, this may result in biased causal inferences (e.g., prior cooperation and the risk of reverse causality would be an example of a potential unmeasured confounder). The risk of potential unmeasured confounding will be discussed with the research team on a site-specific basis; if additional potential confounders are identified and have been measured they will be included in these models; while if additional potential confounders are identified but have *not* been measured this will be discussed when interpreting our site-specific results.

Other features of this DAG are worth noting:

(i) Game features, such as comprehension checks and game order, are included as a node in this DAG. Technically these are not confounders (as they do not cause the exposure), but their inclusion from a causal perspective is benign and may help improve precision in our estimates.

(ii) This DAG assumes no measurement error. That is, all variables are measured perfectly and without error. While this a strong assumption and may be unrealistic in practice - with measurement error in our exposures, outcomes and/or confounders potentially leading to bias and imprecision in our estimates (Greenland, 1980; van Smeden et al., 2020) - our measures are theoretically informed and designed to capture our concepts as closely as possible, hopefully minimising the risk of such bias.

(iii) We also assume no selection bias (Hernán and Robins, 2020). We aim to alleviate this risk by obtaining broadly representative samples of our populations (i.e., random samples based on census data, or comparing sample demographics to population statistics if non-random sampling methods are used). Missing data can also result in selection bias if participation is dependent on the exposure and/or outcome (Hughes et al., 2019). Based on our previous experience with these methods we expect missing data to be minimal, hence unlikely to result in bias, although such missing data may result in reduced statistical power. We will explore missing data on a case-by-case basis and will not discuss it further here, but techniques such as imputing missing data will be used if deemed appropriate.

In sum, reliable causal inference - especially from cross-sectional data - is extremely difficult to achieve and rests on many assumptions which are often impossible to verify empirically. However, as our theories make causal claims, we want to try and make causal inferences using our data and analyses. This is especially pertinent given the difficulties (if not impossibility) of experimentally manipulating supernatural religious beliefs, and the current lack of longitudinal data to help make stronger causal inferences (which would help overcome many - but not all - of the inferential issues discussed above; Bulbulia, 2024).

4 Data Collection

Here we outline the broad data collection methods used across these field sites. These protocols were adapted from the methods of the previous ERM data collections (Lang et al., 2024; Purzycki et al., 2024). Note that here we only provide a broad overview of the data (see the protocols and final reports for more details and site-specific data collection methods). Based on previous power analyses for being able to reliably estimate differences in cooperation in the Random Allocation Game (RAG; see below), we aimed for a minimum of 80 participants per site (for more details of this power analysis, see the grant proposal in the project repository). All materials were first translated into the local language, and subsequently back-translated and edited for consistency and clarity. All of the original protocol materials in English and in translation are or will be available publicly at the repository (`https://github.com/bgpurzycki/GGSL-Project`). Any deviations from this core methodology - such as adjustments to the local context - will be described in the site-specific papers.

Data collection consisted of four modules:

1. Religious Landscape Interview: This was an initial ethnographic interview with approximately 20 participants to understand the local religious landscape. Participants were asked to free-list (Purzycki, 2025) all relevant gods they could think of, then rate whether each god: i) was concerned about people's behaviour; ii) could see into people's minds and their thoughts and feelings; iii) ever punished people for their behaviour; and iv) rewards people for their behaviour. From these questions, two deities were chosen for the main data collection, a 'moralistic god' (MG) who is more punitive and knowledgeable, and a 'local god' (LG) who is relatively less punitive and knowledgeable. If field-workers had previous detailed knowledge of the religious landscape of their site and/or knew the relevant MG and LG in advance, this data was not collected.

2. Demographic survey: The demographic survey was a short questionnaire of 26 items assessing general sociodemographic characteristics. Of particular relevance here are key variables used as core confounders: age, sex, education, number of children and material security.

3. Religiosity survey: To measure religiosity, we asked a battery of 62 questions regarding individuals' beliefs and practices towards the chosen MG and LG at each site. Here, we focus on the questions relating to beliefs

about the LG. Two questions regarding supernatural punishment were asked: "Does [LG] ever punish people for their behaviour?" and "Can [LG] influence what happens to people after they die?". Two questions were also asked regarding omniscience: "Can [LG] see into people's hearts or know their thoughts and feelings?" and "Can [LG] see what people are doing if they are far away in XXXX [a distant town/city]?". All questions were binary yes/no variables. For both punishment and omniscience, as per previous work in this area both questions for each mini-scale will be averaged together for these analyses (Purzycki et al., 2016; Lang et al., 2019). This is for consistency with previous work, and for both practical reasons (to limit the number of interaction terms; see below), and for theoretical reasons (as each question for each mini-scale likely taps into the same/similar construct).

4. We assessed cooperation using two experimental games: the Random Allocation Game (RAG: Hruschka et al., 2014) and the Dictator Game (DG). The RAGs were designed to measure impartial rule-following, while the DGs assess generosity and cooperation more broadly. Each game consisted of two variants with two dyadic recipients. In one, participants divided money between themselves and a geographically distant co-religionist of the same local religious tradition. In the other, participants divided money between local and distant co-religionists of the same local religious tradition. Participants played the RAGs first, followed by the DGs, with the order of recipient counterbalanced within each game. For the RAG, participants were given two cups with lids, one for each recipient (with slots to put coins in). They were then asked to select, in their mind, which cup to put a coin into, then roll a two-colored die (which was previously checked for fairness). Participants were supposed to put a coin into the cup they thought of if the die came up black, else they were supposed to put the coin into the other cup if the die came up white. Participants played the games alone, hence given the anonymous conditions they could choose to ignore the rules and allocate coins to whichever cup they preferred. Each RAG consisted of 30 decisions and corresponding rolls of the die. Systematic, statistically detectable deviations from a 50% chance indicate biasing allocations. For each DG, participants decided how to divide ten allocations of money between the relevant recipients. Each RAG and DG allocation was equivalent to approximately 2% of a days wage, with total earnings given to participants and recipients after the games had been played. Participants were also given a show-up fee of approximately 25% of a day's wage of the local economy.

5 Statistical Models

Our statistical models will follow directly from our causal models, using the data discussed above. While these models will be adapted across the individual site-specific papers, for the purposes of this document we assume the above causal model to be true, and will include all confounders (age, sex, education, material security and number of children) and game features (comprehension checks and game order) as covariates in core regression models; based on the above causal assumptions, this will return an unbiased causal effect. Below, we outline the analysis strategy we anticipate the majority of site-specific papers will follow, while noting that deviations from their broad outline are to be expected.

Analyses will be performed in the statistical software R (Team, 2024), with Bayesian regression models conducted using Stan (Team, 2026) in the package 'brms' (Bürkner, 2017). We will model RAG outcomes with aggregated binomial models, and model DG outcomes with ordinal models. For both games, separate models will be performed for both conditions (self vs. distant co-religionist and local co-religionist vs. distant co-religionist), with "number of coins to distant co-religionists" as the outcome. For all models, we will use weakly-informative priors, with four chains of 2,000 iterations (1,000 as warm-up). Model convergence and sufficient posterior sampling will be assessed via R-hat values (close to 1.00), effective sample sizes and visual inspection of chain convergence.

Here is the structure of the core statistical model for RAGs using an aggregated binomial model:

$$
\begin{aligned}
y_i &\sim \text{Binomial}(30, p_i) \\
\text{logit}(p_i) &= \alpha + \beta_{pun} * \text{pun}_i + \beta_{omni} * \text{omni}_i \\
&\quad + \beta_{age} * \text{age}_i + \beta_{sex} * \text{sex}_i + \beta_{edu} * \text{edu}_i \\
&\quad + \beta_{sec} * \text{sec}_i + \beta_{kids} * \text{kids}_i \\
&\quad + \beta_{order} * \text{order}_i + \beta_{comp} * \text{comprehension}_i \\
\alpha &\sim \text{Normal}(0, 2) \\
\beta &\sim \text{Normal}(0, 1)
\end{aligned}
$$

And here is the equivalent structure of the ordered categorical model for the DGs (note that as there are $\kappa = 11$ possible responses - from 0 to 10 allocations - there are κ - 1 = 10 intercept cut-points in this model):

$$
\begin{aligned}
y_i &\sim \text{Ordered logit}(\phi_i, \kappa) \\
\phi_i &= \beta_{pun} * \text{pun}_i + \beta_{omni} * \text{omni}_i \\
&\quad + \beta_{age} * \text{age}_i + \beta_{sex} * \text{sex}_i + \beta_{edu} * \text{edu}_i \\
&\quad + \beta_{sec} * \text{sec}_i + \beta_{kids} * \text{kids}_i \\
&\quad + \beta_{order} * \text{order}_i + \beta_{comp} * \text{comprehension}_i \\
\beta &\sim \text{Normal}(0, 1) \\
\kappa_k &\sim \text{Normal}(0, 2)
\end{aligned}
$$

To facilitate interpretation of effect sizes on a meaningful scale, we will use g-computation methods to estimate the marginal causal effect of a change in the exposures - from ‘punishment’ and ‘omniscience’ coded as ‘0’ to both coded as ‘1’ - on the outcome, holding all other covariates constant (Hernán and Robins, 2020). These marginal causal effects will denote the average change in the outcome on the count scale (i.e., “an X increase in donations to distant co-religionists”), based on the posterior predictions from the models. All models will be repeated assuming both additive exposure effects (i.e., the effects of ‘punishment’ and ‘omniscience’ on the outcome are linear) and multiplicative exposure effects (i.e., the effects of ‘punishment’ and ‘omniscience’ on the outcome are non-linear, through the use of an interaction term between these exposures). LOOIC information criteria will be used to determine whether the multiplicative model is a better fit to the data than an additive model (although we note that our analyses may be underpowered to detect such interaction effects, given the relatively small sample sizes).

Example analysis code for this core causal model is available here in this project folder (`https://github.com/bgpurzycki/GGSL-Project`). Note that alternative model specifications will be used in the site-specific papers as appropriate based on consideration of the local ethnographic context in collaboration with the field researchers and on the state of the raw data (e.g., inclusion of additional confounders, exploring different assumptions regarding confounding vs mediation for some variables, imputation approaches for missing data, etc.).

Note that, as of February 4, 2026, data for some sites in this project have already been collected and audited (although data from some other sites still needs to be collected and/or audited). We therefore have some information about the data from some of these sites, which has been useful to help understand the complexities of the data and help inform analysis decisions (e.g., patterns and extent of missing data, questions participants had difficulties answering, etc.). However, despite having access to some this data, we assert that prior to finalizing this causal model we did not conduct any preliminary analyses on these data. Our research questions of interest and the structure of our core causal model are therefore independent of any knowledge about the relations between key variables of interest across all GGSL sites (for full transparency, we note that descriptive statistics for key variables and regressions using the core causal model for four sites - Altai Uriankhai, Buryat, Kichwa and Sihanaka - were conducted in January 2026 for a project workshop, but this was *after* the core causal model outlined here had been finalized).

Acknowledgments

This project was financially supported by a grant from the Templeton Religion Trust (#TRT-2022-31107).

References

Baimel, A., Apicella, C., Atkinson, Q., Bolyanatz, A., Cohen, E., Handley, C., Henrich, J., Kundtová Klocová, E., Lang, M., Lesogorol, C., Mathew, S., McNamara, R., Moya, C., Norenzayan, A., Placek, C. D., Soler, M., Vardy, T., Weigel, J., Willard, A., Xygalatas, D., and Purzycki, B. (2022). Material insecurity predicts greater commitment to moralistic and less commitment to local deities: a cross-cultural investigation. *Religion, Brain & Behavior*, 12(1-2):4–17.

Bulbulia, J. A. (2024). Methods in causal inference. Part 1: causal diagrams and confounding. *Evolutionary Human Sciences*, 6:e40.

Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, 80:1–28.

Greenland, S. (1980). The effect of misclassification in the prescence of covariates. *American Journal of Epidemiology*, 112(4):564–569.

Hernán, M. A. and Robins, J. (2020). *Causal Inference: What If.* Chapman & Hall/CRC, Boca Raton.

Hruschka, D., Efferson, C., Jiang, T., Falletta-Cowden, A., Sigurdsson, S., McNamara, R., Sands, M., Munira, S., Slingerland, E., and Henrich, J. (2014). Impartial Institutions, Pathogen Stress and the Expanding Social Network. *Human Nature*, 25(4):567–579.

Hughes, R. A., Heron, J., Sterne, J. A. C., and Tilling, K. (2019). Accounting for missing data in statistical analyses: multiple imputation is not always the answer. *International Journal of Epidemiology*, 48(4):1294–1304.

Johnson, D. and Krüger, O. (2004). The Good of Wrath: Supernatural Punishment and the Evolution of Cooperation. *Political Theology*, 5(2):159–176. Publisher: Routledge _eprint: https://doi.org/10.1558/poth.2004.5.2.159.

Johnson, D. D. P. (2005). God's punishment and public goods. *Human Nature*, 16(4):410–446.

Lang, M., Purzycki, B. G., Apicella, C. L., Atkinson, Q. D., Bolyanatz, A., Cohen, E., Handley, C., Kundtová Klocová, E., Lesorogol, C., Mathew, S., McNamara, R. A., Moya, C., Placek, C. D., Soler, M., Vardy, T., Weigel, J. L., Willard, A. K., Xygalatas, D., Norenzayan, A., and Henrich, J. (2019). Moralizing gods, impartiality and religious parochialism across 15 societies. *Proceedings of the Royal Society B: Biological Sciences*, 286(1898):20190202.

Lang, M., Purzycki, B. G., Henrich, J., and Norenzayan, A. (2024). *The Evolution of Religion and Morality: Volume II.* Routledge.

Lundberg, I., Johnson, R., and Stewart, B. M. (2021). What Is Your Estimand? Defining the Target Quantity Connects Statistical Evidence to Theory. *American Sociological Review*, 86(3):532–565. Publisher: SAGE Publications Inc.

Major-Smith, D. (2023). Exploring causality from observational data: An example assessing whether religiosity promotes cooperation. *Evolutionary Human Sciences*, 5:e22.

McElreath, R. (2020). *Statistical Rethinking: A Bayesian Course with Examples in R and STAN*. Chapman & Hall/CRC, 2nd edition.

Norenzayan, A., Shariff, A. F., Gervais, W. M., Willard, A. K., McNamara, R. A., Slingerland, E., and Henrich, J. (2016). The cultural evolution of prosocial religions. *Behavioral and Brain Sciences*, 39:e1.

Pearl, J., Glymour, M., and Jewell, N. (2016). *Causal Inference in Statistics: A Primer*. Wiley.

Purzycki, B. G. (2025). *Ethnographic free-list data: Management and analysis with examples in R*. Sage.

Purzycki, B. G., Apicella, C., Atkinson, Q. D., Cohen, E., McNamara, R. A., Willard, A. K., Xygalatas, D., Norenzayan, A., and Henrich, J. (2016). Moralistic gods, supernatural punishment and the expansion of human sociality. *Nature*, 530(7590):327–330.

Purzycki, B. G. and Bendixen, T. (2024). Material insecurity and religiosity: A causal analysis. *Evolutionary Human Sciences*, 6:e4.

Purzycki, B. G., Henrich, J., and Norenzayan, A. (2024). *The Evolution of Religion and Morality: Volume I*. Routledge.

Purzycki, B. G., Lang, M., Henrich, J., and Norenzayan, A. (2022). The Evolution of Religion and Morality project: reflections and looking ahead. *Religion, Brain & Behavior*, 12(1-2):190–211.

Purzycki, B. G. and Sosis, R. (2022). *Religion Evolving: Cultural, Cognitive and Ecological Dynamics*. Equinox, Bristol, CT.

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. Publisher: SAGE Publications Inc.

Schloss, J. P. and Murray, M. J. (2011). Evolutionary accounts of belief in supernatural punishment: a critical review. *Religion, Brain & Behavior*, 1(1):46–99. Publisher: Routledge _eprint: https://doi.org/10.1080/2153599X.2011.558707.

Team, R. C. (2024). R: A Language and Environment for Statistical Computing.

Team, S. D. (2026). Stan Reference Manual.

van Smeden, M., Lash, T. L., and Groenwold, R. H. H. (2020). Reflection on modern methods: five myths about measurement error in epidemiological research. *International Journal of Epidemiology*, 49(1):338–347.

Vardy, T., Moya, C., Placek, C. D., Apicella, C. L., Bolyanatz, A., Cohen, E., Handley, C., Kundtová Klocová, E., Lesorogol, C., Mathew, S., McNamara, S. A., Purzycki, B. G., Soler, M., Weigel, J. L., Willard, A. K., Xygalatas, D., Norenzayan, A., Henrich, J., Lang, M., and Atkinson, Q. D. (2022). The religiosity gender gap in 14 diverse societies. *Religion, Brain & Behavior*, 12(1-2):18–37.