Preregistration

Explaining Atheism: Pilot 1 Preregistration

Connair J. S. Russell[1], Jon Lanman[1], Aiyana Willard[2], Adam Baimel[3]

[1] Queen's University Belfast

[2] Brunel University London

[3] Oxford Brookes University

INSERT DATE

Study Information

Title Explaining Atheism: Pilot 1 Preregistration

Description The present work is part of the broader Explaining Atheism project, which looks to examine popular and scientific causes of belief/non-belief and how these vary across cultures.

Across 3 stages we will examine variables which have been argued in popular and scientific discourse to cause belief and non-belief. In this initial pilot we will look at the largest number of variables in order to test some of the clearest hypotheses and to run explanatory tests to exclude possible variables for the later stages of the project which will have larger samples and a greater proportion of confirmatory tests.

This pre-registration is for analysis of our first pilot data, with analyses looking to include or exclude variables for later stages. Here we look at the most popular

explanations of belief/non-belief and whether they predict a number of different measures related to belief/non-belief

Additional preregistrations for this stage testing specific hypotheses, along with preregiistrations for follow on pilots, can be found on our pages on the Open Science Framework, our Figshare, and on my personal GitHub.

Hypotheses

The current pre-registration is for exploratory analyses, as such no explicit directional hypotheses are tested. We will instead be testing the hypotheses of our explanatory variables having an effect on one or more outcome variables.

Design Plan

Study type

Observational (Survey) Study

Blinding

No blinding is involved in this study.

Study design

This study uses a survey design.

This pilot will run 5 separate surveys, one for each of the explanatory variable clusters, in the UK, Japan and Brazil. Survey 5 will also be run in Denmark. The results in 16 independent survey samples.

Randomization

Question items will be randomised *within* surveys, with the following exceptions:

-
-
-

Sampling Plan

We will collect representative samples for each survey for each nation. This will be done by an independent survey company, and the samples will be representative for the national population by Age, Sex and Geographic Region.

Existing data	**Registration prior to creation of data**. As of the date of submission of this preregistration, the data have not yet been collected, created, or analysed.
Data collection procedures	Participants are recruited by the independent survey company, Savanta Comres, from their pool of participants. Participants will be paid a local rate for their participation, and will be selected to meet representativeness requirements outlined above. ***Exclusion criteria***
Sample size	The sample size used is 500 per survey per nation, resulting in samples per nation of: Brazil = 2500 UK = 2500 Japan = 2500 Denmark = 500
Sample size rationale	Samples are the maximally viable sample for the financial means of the project.

Variables

Please find below a list of each variable planned to be measured in the study. A background for each of these, along with their operationalisations, subscales, and scripts can be found in the codebook accompanying the pre-registration.

Explanatory Variables

Morals/ Values

- Individualism & Collectivism
- Rebelliousness (complete)
- Rebelliousness (trolling)
- Rebelliousness (activism)
- Individual Choice Norms

Cognitive Style

- Tolerance of Ambiguity
- Cognitive Style (correct)
- Cognitive Style (intuitive)
- Non-Verbal Reasoning
- Moralisation of Rationality
- Importance of Rationality

Socialisation

- CRED Exposure
- CRED (caregiver)
- CRED (Group Leader)
- CRED (Group Member)
- CRED (Group)
- CRUD Exposure
- CRUD (Caregiver)

- CRUD (Group Leaders)
- CRUD (Group)
- Non-theistic Socialisation
- Normativity of Religion (All)
- Normativity of Religion (Nation)
- Normativity of Religion (Ethnicity)
- Normativity of Religion (Peers)
- Normativity of Religion (Group)
- Religious Emphasis
- Parental Circumstances
- Parental Absence Reason

Motivational

- Social Desirability
- Existential Security (Current)
- Existential Security (Future)
- Existential Security (Upbringing)
- Social Security
- Wealth Inequality
- Social Trust
- Need for Structure
- Death Anxiety
- Social Network Size
- Meaning in Life
- Need for Meaning
- Disgust Sensitivity
- Disgust Sensitivity (Core)
- Disgust Sensitivity (Animal)
- Disgust Sensitivity (Contamination)

Cognitive Bias

- Anthropomorphism

- Mentalizing
- Dualism
- Pattern Perception
- Pattern Perception (Long)
- Promiscuous Teleology
- Schizotypy
- Schizotypy (Reference)
- Schizotypy (Perceptual Experiences)
- Vividness of Mental Imagery
- Dissociative Absorption

Outcome Variables

Belief in God

- Belief in God (Categorical)
- Belief in God (Continuous/Ordinal)
- Belief in God (Binary)
- Agnosticism (Continuous/ Ordinal)
- Agnostic Identity (binary)

Atheism

- Atheist Membership
- Atheist Identity

Religiosity

- Religious Identification
- Religious Identity
- Religious Attendance
- Prayer Frequency
- Religious Objects
- Anti-religiosity

Meta Belief

- Possibility of Knowing (God)
- Confidence in Belief
- Apatheism (God)
- Apatheism (Life Purpose)
- Apatheism (Combined)
- Possibility of Truth
- Naturalism

Supernatural Belief

- Afterlife Existence
- Afterlife Punishment
- Afterlife Reward
- Reincarnation
- Astrology
- Mystical People
- Mystical Objects
- Good and Evil
- Universal Spirit or Life Force
- Karma
- Fate
- Supernatural Being Existence
- Good Supernatural Beings
- Harmful Supernatural Beings
- Evil Eye
- Spiuritual Experience (Personal)
- Lucky Object

Indices Coding and scoring methods are included for each variable in the accompanying codebook. Draft scoring scripts are also included.

Analysis Plan

Exploratory Analysis Plan

We will look at the effects of our explanatory variables on various outcome variables. However, in order to examine which variables should be included in later rounds of our research, we will primarily look at the effects of our different clusters of explanatory variables (***bold***) on the following outcome variables:

Morals/ Values:

Belief in God; Religious Identity, Atheist Identity, Agnosticism; Religious Attendance

Cognitive Styles:

Belief in God; Religious Identity, Atheist Identity, Agnosticism; Religious Attendance

Socialisation

Belief in God; Religious Identity, Atheist Identity, Agnosticism; Religious Attendance

Cognitive Biases

Belief in God; Religious Identity, Atheist Identity, Agnosticism; Religious Attendance; Supernatural Being Existence

Motivational

Belief in God; Religious Identity, Atheist Identity, Agnosticism; Religious Attendance

---

Statistical models

To examine the possible relationships between our explanatory and our outcome variables we will run bayesian multiple regression models in R using the '{brms}' package. We will run models for all the outcome and explanatory variable combinations noted in the above analysis plan, among others.

Model specifications can be found in the attached Excel files.

We will treat outcome variables that are single Likert scale measures as ordinal, and as such will use cumulative probit regression following the following structure:

```r
library("brms")



exampleProbitModel <-
  brm(formula = outcome_variable ~ explanatory_variable_1 +
                                   explanatory_variable_2 +
                                   explanatory_variable_3 +
                                   explanatory_variable_n,
      prior = priors,
      data = data,
      family = cumulative(link = "probit"),
      chains = 4,
      cores = 4,
      threads = 2,
      iter = 3000,
      warmup = 1000,
      seed = 2023,
      backend = "cmdstanr")
```

For binary outcomes we will use logistic regression with the following format:

```r
library("brms")


exampleLogisticModel<-
  brm(formula = outcome_variable ~ explanatory_variable_1 +
                                   explanatory_variable_2 +
                                   explanatory_variable_3 +
                                   explanatory_variable_n,
      prior = priors,
      data = data,
      family = bernoulli(link = "logit"),
      chains = 4,
      cores = 4,
      threads = 2,
```

```
      iter = 3000,
      warmup = 1000,
      seed = 2023,
      backend = "cmdstanr")
```

And for continuous outcomes we will use gaussian multiple regression:

```
library("brms")


exampleLogisticModel<-
  brm(formula = outcome_variable ~ explanatory_variable_1 +
                                   explanatory_variable_2 +
                                   explanatory_variable_3 +
                                   explanatory_variable_n,
      prior = priors,
      data = data,
      family = gaussian(),
      chains = 4,
      cores = 4,
      threads = 2,
      iter = 3000,
      warmup = 1000,
      seed = 2023,
      backend = "cmdstanr")
```

Priors

We will set the following uninformative priors for our outcome variables:

For betas and intercepts we will set a normally distributed prior with a mean of 0 and a standard deviation of 1.

For ordinal probit models we will set a normal distribution for the thresholds, following Kurz.

Transformations	In some instances response variables are dummy coded binary variables from a categorical response option - see attached codebook for all variable definitions. As a default we will center likert response scores at the center of the scale, unless the scale has a meaningful interval (e.g. age, n years).
Inference criteria	In the first instance, inferences about whether a variable has or does not have a theoretically relevant effect will be done based on examination of the posterior distributions of parameter estimates. More specifically we will examine the extent to which the credible intervals of these posterior distributions overlap with a Region of Practical Equivalence (ROPE), that is, a specified region we consider to be theoretically equivalent to zero. This will be done using the {bayestestR} package (https://easystats.github.io/bayestestR/articles/region_of_practical_equivalence.html) in R. For our Gaussian and Cumulative Probit models we will use the default range -0.1 to 0.1 for standarized effect sizes, which equates to a "negligible effect size", and -0.18 - 0.18 for logistic models.
Data exclusion	Participants who fail data quality checks are removed and replaced by the survey provider prior to our analysis. In the first instance outliers will not be removed, however sensitivity analysis will be run on any published final models.
Missing data	No missing data will be in
Exploratory analyses (optional)	See Exploratory Data Analysis Plan above.

Other

Other (Optional)	Enter your response here.

References