Super-basic results overview

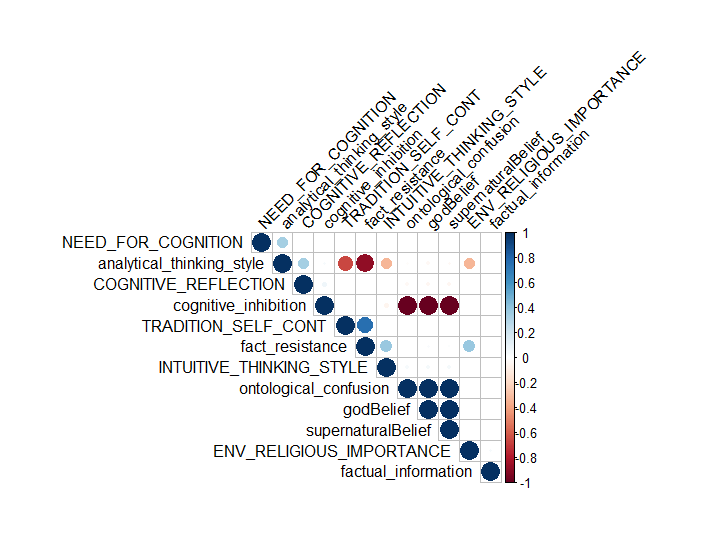
# Correlation and Regression

To better understand the model under all possible conditions, we ran a “parameter sweep” where we ran the model under 20,000 unique conditions where we vary each of our input parameters between their minimum and maximum values. This gives us a full sweeping of the space so that we can better understand the different conditions of the model. This is useful in some ways, but let’s keep in mind how, given the range of all possible conditions, how many of them might be applicable to understanding society as it operates today. That is to say, there may be cases that are theoretically possible in our model that aren’t interesting (or maybe feasible) in the real world.

Nevertheless, this lets us better understand if our model is capturing relationships that are observed, generally speaking, in the empirical literature. The two most reasonable ways of doing this is through correlation and regression—much as it is in the empirical literature.

## Correlation

We can start by taking the variables that we gathered from our parameter sweep of the model and looking for significant correlations between them. In this case, we just correlated all variables against the others to see what relationships there are in the model. These are visualized in the plot below. Interactions that did not achieve significance do not have any coloration or markings in the matrix below (these are all Spearman’s rho due to non-normality resulting from the Latin Hypercube Sampling technique).



## Regression

We also utilized regression to understand the positive and negative effects of different variables on god beliefs and supernatural beliefs.

The output from regressing all beliefs onto our “godBelief” variable (representing religious beliefs) are as such:

Residuals:

Min 1Q Median 3Q Max

-8.149e-13 -1.180e-15 -7.000e-17 1.050e-15 8.138e-13

Coefficients: (3 not defined because of singularities)

Estimate Std. Error t value Pr(>|t|)

(Intercept) -2.780e-15 1.545e-14 -1.800e-01 0.8572

PRESENCE\_OF\_A -1.828e-15 4.510e-16 -4.053e+00 5.08e-05 \*\*\*

FIRST\_TRAUMA -2.453e-18 1.576e-19 -1.556e+01 < 2e-16 \*\*\*

FREQ\_OF\_TRAUMA -7.393e-19 1.111e-19 -6.654e+00 2.93e-11 \*\*\*

INTELLIGENCE -1.000e+00 4.507e-16 -2.219e+15 < 2e-16 \*\*\*

INIT\_COG\_INHIBITION 3.138e-15 4.506e-16 6.964e+00 3.41e-12 \*\*\*

COG\_INHIB\_DEPLETION 1.258e-14 4.532e-16 2.775e+01 < 2e-16 \*\*\*

INIT\_RELIGIOUS\_INFO 3.139e-15 4.507e-16 6.965e+00 3.38e-12 \*\*\*

FAMILY\_RELIGIOUS\_IMPORTANCE 2.000e+00 4.508e-16 4.436e+15 < 2e-16 \*\*\*

INIT\_FACT\_INFO 1.054e-15 4.766e-16 2.212e+00 0.0270 \*

EVENT\_INTENSITY -2.348e-15 4.506e-16 -5.211e+00 1.89e-07 \*\*\*

START\_UG\_AGE 1.315e-17 1.865e-18 7.050e+00 1.85e-12 \*\*\*

START\_PG\_AGE -2.916e-18 1.358e-18 -2.147e+00 0.0318 \*

END\_UG\_AGE NA NA NA NA

STUDYTOPIC 9.796e-16 4.509e-16 2.173e+00 0.0298 \*

TRADITION\_SELF\_CONT 2.000e+00 2.253e-16 8.877e+15 < 2e-16 \*\*\*

INTUITIVE\_THINKING\_STYLE 2.000e+00 4.599e-16 4.349e+15 < 2e-16 \*\*\*

COGNITIVE\_REFLECTION -1.000e+00 4.626e-16 -2.162e+15 < 2e-16 \*\*\*

NEED\_FOR\_COGNITION -1.000e+00 4.506e-16 -2.219e+15 < 2e-16 \*\*\*

ENV\_RELIGIOUS\_IMPORTANCE 1.000e+00 4.508e-16 2.218e+15 < 2e-16 \*\*\*

fact\_resistance NA NA NA NA

cognitive\_inhibition -1.049e-15 1.711e-17 -6.133e+01 < 2e-16 \*\*\*

factual\_information 2.236e-17 1.543e-16 1.450e-01 0.8848

ontological\_confusion 1.000e+00 1.720e-17 5.813e+16 < 2e-16 \*\*\*

analytical\_thinking\_style NA NA NA NA

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.839e-14 on 19978 degrees of freedom

Multiple R-squared: 1, Adjusted R-squared: 1

F-statistic: 1.524e+34 on 21 and 19978 DF, p-value: < 2.2e-16

The output from regressing all beliefs onto our “supernaturalBelief” variable (representing spiritual—but not religious—beliefs) are as such:

Residuals:

Min 1Q Median 3Q Max

-9.48e-13 -1.11e-15 -2.00e-17 1.07e-15 8.82e-13

Coefficients: (3 not defined because of singularities)

Estimate Std. Error t value Pr(>|t|)

(Intercept) -7.422e-15 1.491e-14 -4.980e-01 0.619

PRESENCE\_OF\_A -4.584e-15 4.353e-16 -1.053e+01 < 2e-16 \*\*\*

FIRST\_TRAUMA -8.134e-19 1.521e-19 -5.348e+00 9.00e-08 \*\*\*

FREQ\_OF\_TRAUMA -1.101e-18 1.072e-19 -1.027e+01 < 2e-16 \*\*\*

INTELLIGENCE -1.000e+00 4.349e-16 -2.299e+15 < 2e-16 \*\*\*

INIT\_COG\_INHIBITION -2.491e-15 4.349e-16 -5.728e+00 1.03e-08 \*\*\*

COG\_INHIB\_DEPLETION 7.411e-15 4.373e-16 1.695e+01 < 2e-16 \*\*\*

INIT\_RELIGIOUS\_INFO 5.062e-15 4.349e-16 1.164e+01 < 2e-16 \*\*\*

FAMILY\_RELIGIOUS\_IMPORTANCE 1.000e+00 4.351e-16 2.298e+15 < 2e-16 \*\*\*

INIT\_FACT\_INFO -2.840e-15 4.599e-16 -6.174e+00 6.77e-10 \*\*\*

EVENT\_INTENSITY 2.298e-16 4.349e-16 5.280e-01 0.597

START\_UG\_AGE 5.605e-19 1.800e-18 3.110e-01 0.756

START\_PG\_AGE -2.036e-18 1.311e-18 -1.553e+00 0.120

END\_UG\_AGE NA NA NA NA

STUDYTOPIC 8.863e-15 4.351e-16 2.037e+01 < 2e-16 \*\*\*

TRADITION\_SELF\_CONT 1.000e+00 2.174e-16 4.599e+15 < 2e-16 \*\*\*

INTUITIVE\_THINKING\_STYLE 2.000e+00 4.438e-16 4.506e+15 < 2e-16 \*\*\*

COGNITIVE\_REFLECTION -1.000e+00 4.464e-16 -2.240e+15 < 2e-16 \*\*\*

NEED\_FOR\_COGNITION -1.000e+00 4.349e-16 -2.299e+15 < 2e-16 \*\*\*

ENV\_RELIGIOUS\_IMPORTANCE 1.000e+00 4.351e-16 2.298e+15 < 2e-16 \*\*\*

fact\_resistance NA NA NA NA

cognitive\_inhibition -1.049e-15 1.651e-17 -6.355e+01 < 2e-16 \*\*\*

factual\_information 2.236e-17 1.490e-16 1.500e-01 0.881

ontological\_confusion 1.000e+00 1.660e-17 6.024e+16 < 2e-16 \*\*\*

analytical\_thinking\_style NA NA NA NA

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.775e-14 on 19978 degrees of freedom

Multiple R-squared: 1, Adjusted R-squared: 1

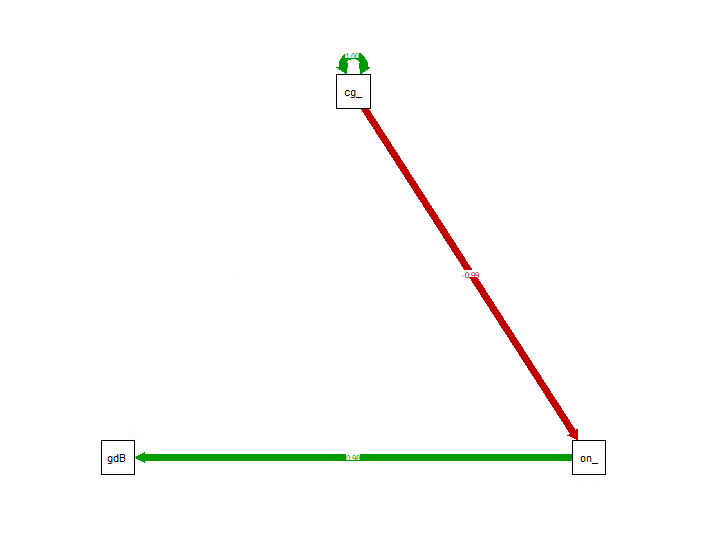
F-statistic: 1.636e+34 on 21 and 19978 DF, p-value: < 2.2e-16

In order to better understand the key variables of our model, we can trim down to just the most important variables from the literature. This regression results are found below.

|  |  |  |  |
| --- | --- | --- | --- |
| Variable | Beta (standardized) | t | p |
| INTELLIGENCE | -0.0039 | -199.49 | <.01 |
| TRAD\_SELF\_CONT | 0.0039 | 115.22 | <.01 |
| INTUITIVE\_THINKING\_STYLE | 0.0019 | 79.30 | <.01 |
| COGNITIVE\_REFLECTION | -0.0039 | -194.13 | <.01 |
| NEED\_FOR\_COGNITION | -0.0039 | -199.77 | <.01 |
| Fact\_resistance | 0.016 | 421.99 | <.01 |
| Cognitive\_inhibition | 0.00034 | 1.80 | .07 |
| Ontological\_confusion | 0.99 | 5243.04 | <.01 |

General model statistics: *RSE* = 2.05, *df* = 19991, *r2* = 1, *f* = .000000032(8,19991), *p* < .01.

We were somewhat surprised to see the insignificant relationship with cognitive\_inhibition in the data. To further investigate our system we used the Lavaan package to create an SEM of a mediation model to test that the relationship between ontological confusion and god beliefs was mediated by cognitive inhibition. Indeed, we found that the relationship between cognitive inhibition and god beliefs is mediated by ontological confusion and that this relationship can be expressed as a full mediation model.



# Optimization

As with before I used AnyLogic to run optimization experiments. In doing this, we ask the computer to find us the conditions under which something is maximized or minimized. It does this by varying our key inputs (parameters; e.g., need for cognition, intuitive thinking style, tradition and self continuity, and cognitive inhibition) and finding the set of parameters that get us closest to the value that we’re optimizing.

We asked the computer to assess “godBeliefs” and “spiritualBeliefs” (generally interpreted as religious beliefs and spirituality respectively). We asked the simulation to find 5 conditions: 1) find the parameters where religious beliefs is highest, and spirituality is lowest (i.e., “I’m religious but not spiritual”); 2) find the parameters where religious beliefs is minimized and spirituality is maximized (i.e., “I’m spiritual but not religious”); 3) find the parameters where religious beliefs AND spirituality are both maximized (i.e., “I’m super religious and spiritual); 4) find the parameters where religious beliefs and spirituality are both minimized (i.e., “I’m an avowed strong atheist”); 5) find the parameters where both religious beliefs and spirituality are closest to 0 (i.e., “I’m agnostic”).

The computer found that the variables (listed in the rows) for each of the 5 settings (in the columns) are as such (in the table below). The coloring is there from green (low) to red (high) showing the relative values BETWEEN CONDITIONS (not between variables). This is because the numerical values themselves are not as important as the relative values between conditions. Also, these values do not exist in isolation from one another. So we can’t read this as we would a regression for example. All of the variables are—basically—one part of a single “factor” that results in the condition.



# Calibration

Similarly, we also were able to run calibration experiments. These are like optimization experiments, but instead of trying to find a single output value, we look at a longitudinal time slice, and try and find the parameters that produce output that best matches the desired output. This is done by inputting an idealized step-function into the model that represents what we’re hoping to simulate. So, for example, when trying to simulate a “conversion” from “believer” to “atheist” we would input a “Target” dataset for the simulation to match, then the simulation would match that data as best as possible. An example is included below for reference:

To calibrate our model against different religious life histories, we asked the simulation to calibrate to the following:

|  |  |
| --- | --- |
| Setting | Pattern |
| AlwaysAtheist | always at 0' |
| AlwaysStrongAtheist | always at -500' |
| AlwaysReligious | always at 500' |
| Religious2AtheistConv | 500 to 0 at age 20' |
| Religious2AtheistConvLate | 500 to 0 at half way' |
| Atheist2ReligiousConv | 0 to 500 at age 20' |
| Atheist2RelConvLate | 0 to 500 at half way' |
| Atheist2RelConvVERYlate | 0 to 500 at age 60' |
| StrAtheist2ReligiousConv | -500 to 100 at age 20' |
| StrAtheist2RelConvLate | -500 to 100 at half way' |
| StrAtheist2RelConvVERYlate | -500 to 500 at age 60' |

The computer found that the variables (listed in the rows) for each of the 5 settings (in the columns) are as such (in the table below). The coloring is there from green (low) to red (high) showing the relative values BETWEEN CONDITIONS (not between variables). This is because the numerical values themselves are not as important as the relative values between conditions. Also, these values do not exist in isolation from one another. So we can’t read this as we would a regression for example. All of the variables are—basically—one part of a single “factor” that results in the condition.

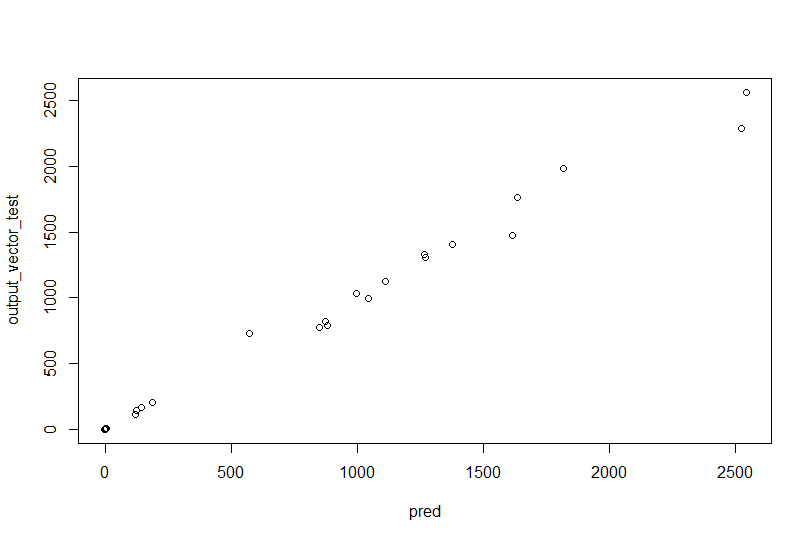


# Key Features

Lastly, we decided to utilize techniques from Artificial Intelligence (AI) to better understand our data. One important aspect of creating AI models using machine learning and statistical learning techniques is identifying the key features from a dataset that are driving some classification problem. We utilized this approach to understanding “big data” on our model’s output.

We started by taking all data and coding it as “believer” or “nonbeliever” depending on if their godBelief value was > 0 or < 0 respectively. We then separated the data into two datasets the first dataset was the first 15,000 simulations, which we used to train the machine learning model, the second was the last 5,000 simulations, which we used to test our machine learning model.

We then created a machine learning model using the XGBoost (extreme gradient boosting) approach, leveraging the r library “xgboost”. We found that our model was able to have a pretty good fit.



From this, we then aimed to find out what the most important factors are for the AI when determining whether or not the input data represents a “believer” or a “nonbeliever”

Our AI system determined that, when classifying a person, based on the input parameters, as a “believer” or “nonbeliever” that the top-5 most important data in determining the classification decision are from the following variables (ranked in order of importance):

Cognitive inhibition

Analytical thinking style

Tradition and self continuity

Frequency of trauma

Presence of atheists

It is worth noting that cognitive inhibition is OVERWHELMINGLY the most important factor that the AI system uses to classify the data.