Brief article

Text analysis shows conceptual overlap as well as domain-specific differences in Christian and secular worldviews

Joseph Watts, Sam Passmore, Joshua Conrad Jackson, Christoph Rzymski, Robin I.M. Dunbar

ARTICLE INFO

Keywords:
Text analytics
Religion
Supernatural belief
Christianity
Secularization
Ontology
World explanations
Open data
Preregistered

ABSTRACT

Theories differ over whether religious and secular worldviews are in competition or represent overlapping and compatible frameworks. Here we test these theories by examining homogeneity and overlap in Christian and non-religious people's explanations of the world. Christian and non-religious participants produced free text explanations of 54 natural and supernatural phenomena. Using a new text analytic approach, we quantitatively measure the similarity between 7613 participant-generated explanations. We find that the relative homogeneity of Christian and non-religious people's explanations vary depending on the kind of phenomena being explained. Non-religious people provided more similar explanations for natural than supernatural phenomena, whereas Christian explanations were relatively similar across both natural and supernatural phenomena. This challenges the idea that religious systems standardize and restrict people's worldviews in general, and instead suggest this effect is domain specific. We also find Christian and non-religious participants used largely overlapping concepts to explain natural and supernatural phenomena. This suggests that religious systems supplement rather than compete with secular based worldviews, and demonstrates how text analytics can help understand the structure of group differences.

1. Introduction

People disagree over what happens when we die, whether we are alone in the universe, and the origins of life on earth. Religions are popularly seen as a major source of this disagreement, but it is not clear exactly how religions affect people's explanations of the world. Do religions force people to use a specific set of ideological concepts to explain the world? Or do religions simply add an additional layer of explanation?

One popular claim is that religions represent systems of concepts, termed “memeplexes,” that replicate at the expense of alternative religious and secular worldviews (Dawkins, 2006; Dennett, 2006). These competition-based accounts predict that the worldviews of people with the same religious affiliation share more concepts than people with different religious affiliations. They also predict that religious systems homogenize adherents' worldviews by prescribing a divinely sanctioned doctrine. In support of competition-based accounts, studies show that priming religious explanations leads people to rely less on scientific explanations (Preston & Epley, 2009) and human agency (Dijksterhuis, Preston, Wegner, & Aarts, 2008).

Alternative theories argue that religious and secular worldviews explain different aspects of the world, and can be combined into a coherent conceptual system (Astuti & Harris, 2008; Gould, 1999; Subbotsky, 2001; Watts, 1997). These synthesis-based accounts predict that religious worldviews include the same kinds of concepts as secular worldviews, with the addition of an expanded set of supernatural concepts. In support of these accounts, research shows that religious individuals simultaneously endorse both natural and supernatural explanations of the same phenomena (Busch, Watson-Jones, & Legare, 2017; Cornelius, Lacy, & Woolley, 2011; Legare, Evans, Rosengren, & Harris, 2012).

Existing empirical studies analyzing the relationship between
natural and supernatural explanations have used Likert-scale responses, experimenter coded binary and categorical classification systems, or reaction times in binary evaluations (Busch et al., 2017; Kelemen, Rottman, & Sexton, 2013; Legare et al., 2012; Nancekivell & Friedman, 2017; Preston & Epley, 2009; Woolley & Cornelius, 2017). These studies have provided important theoretical insights into the cognitive processes underlying supernatural beliefs, but their methodological approaches restrict the amount of information analyzed because they do not represent the extent of semantic overlap in people's explanations. These studies have also focused on how people explain specific kinds of phenomena, such as the causes of illness (Busch et al., 2017; Legare et al., 2012; Preston & Epley, 2009), or the reasons for unexpected, impossible or unlikely events (Cornelius et al., 2011; Nancekivell & Friedman, 2017; Woolley & Cornelius, 2017). One reason that both competition-based and synthesis-based accounts have found empirical support could be that religious people may synthesize science and religious based concepts for some kinds of phenomena, such as the causes of death, but treat religious and scientific concepts as competing for other kinds of phenomena, such as what happens after death.

Here we develop a new way of testing between competition-based and synthesis-based accounts across a variety of domains using text analysis. Text analytic methods can efficiently and systematically quantify the similarity between Christian and non-religious participants' explanations of a diverse range of phenomena. These data allow us to identify whether religious affiliations divide the kinds of concepts that people use to explain the world, and whether religious people use more homogeneous explanations than non-religious people.

2. Methods

2.1. Participants

A total of 245 participants were recruited for this study, but 33 were excluded due to low effort responses or inconsistencies between their pre-screened identity and the post-study questionnaire. For example, we excluded a participant that identified as “Wiccan” in the post-study questionnaire (a full description of exclusions is provided in the Supplementary Materials). After exclusions, our study included 212 participants from the United States of America, 101 who identified as Christians (55 female and 46 male), and 111 who identified as non-religious (51 female and 60 male). The Christian group included non-denominational, Catholic and Protestant Christians, while the non-religious group included participants that identified as agnostic, atheist or having no religion. Participants ranged in age from 18 years to 68 years old (mean = 33.92 years, SD = 10.35 years). The highest qualification of participants was primary school for two participants, high school for 49 participants, college for 45 participants, undergraduate university qualification for 87 participants, and a postgraduate or PhD qualification for 29 participants. Participants were recruited through the Prolific online participant pool, the experimental procedure took approximately an hour to complete, and each participant was reimbursed the equivalent of £9 for their time.

2.2. Study design

The full pre-registered study design, as well as data and reproducible code, are available through the Open Science Framework (OSF) project page (https://osf.io/sgv3h/). In order to address reviewer comments, we doubled the pre-registered sample size of this study. A complete summary of deviations from this pre-registration are available in the Supplementary Materials.

Participants were presented with descriptions of natural and supernatural phenomena and asked to provide what they consider to be the best explanation for each phenomenon. This allowed participants to explain the world in their own words, rather than through fixed scales or multi-choice responses.

There were 54 different descriptions of phenomena in the study, half of which referred to natural phenomena and half of which referred to supernatural phenomena (domain of phenomena). Here, ‘natural’ denotes parts of the Universe that are subject to the laws of nature and amenable to scientific analysis. ‘Supernatural’ denotes those parts of the Universe that are beyond the laws of nature and/or are not amenable to scientific analysis. We balanced the sentiment of phenomena being explained so that there was an even number of negative, neutral, and positive phenomena in each domain. We also split each domain into three further sub-domains, and asked participants to explain an equal number of phenomena for each sub-domain. For the natural domain, the three sub-domains were social, biological and physical phenomena. For the supernatural domain, the three sub-domains were traditional religious belief, superstition, and new age belief. In the Supplementary materials, we provide a full description of how sentiment and sub-domains of phenomena are defined, as well as additional analyses of their effects.

We anticipated that the study would require participants to concentrate for extended periods of time so only presented each participant with 36 of the 54 descriptions of phenomena in order to prevent fatigue. The order that phenomena were presented to participants was randomized. Participants were asked to give a written explanation for each of the 36 phenomena they were presented with. As part of the supplementary study, participants were also asked to specify the extent to which their explanation was natural and/or supernatural. Each phenomenon was presented on a separate page that included a text field and a prompt to provide an explanation. Explanations were constrained to be between 60 and 100 characters, which was enforced through the oTree software package used to run the study (Chen, Schonger, & Wickens, 2016).

At the end of the study, participants filled in a demographic survey, including age, educational level, gender, nationality and religious affiliation. In the post-experiment section participants were also asked to complete the Revised Paranormal Belief Scale (Tobacyk, 2004) which was used as an additional check on participants’ religious and supernatural beliefs. A complete list of our pre-registered variables, models and hypotheses are available on the OSF project page.

2.3. Algorithm for calculating semantic similarity of explanations

We used a text-analytics approach to calculate a continuous measure of conceptual similarity between explanations. This avoided the subjectivity of manual coding, and enabled the quantitative analysis of free-text explanations on a scale not feasible using manual coding. All explanations were cleaned for comparison by removing all punctuation and stop-words (words that don’t contain subject meaning) and converted to lower case. Remaining words were normalized to a common form through a process of lemmatization (Rinker, 2018). For example, drive, drove, and driven are all reduced to the common form drive via a dictionary lookup. This process results in a set of keywords for each explanation that can then be used in comparisons (Tonkin & Tourte, 2016). We used the R text-analytic packages tm and textstem for these processes (Feinerer & Hornik, 2008; Rinker, 2018).

We generated pairwise similarity measures between the keywords of all explanations for the same phenomena. Similarity between pairs of explanations was calculated using Jaccard index, defined as the number of unique overlapping keywords between the two explanations (A ⋃ B), divided by the total number of unique keywords (A ⋃ B) in the two explanations:

\[
\text{Jaccard index} = \frac{|A \cap B|}{|A \cup B|}
\]

where A indicates the set of keywords from one explanation and B indicates the set of words from a second explanation. Only explanations to the same phenomena were compared. If two sets contained exactly the same keywords, they would have a Jaccard similarity of 1. If two
discussed with non-religious participants. We found that Christian's explanations of supernatural phenomena were more similar than between-group explanations (MCMCglmm: posterior mean = –8.31, CrI = –21.90 to 4.48, pMCMC = 0.208). Neither did we find evidence that between-group explanation similarity was greater than the similarity of non-religious people's explanations (MCMCglmm: posterior mean = –1.75, CrI = –1.471 to 12.09, pMCMC = 0.814). While these results do not take into account how the similarity between Christian and non-religious people might vary across the kinds of phenomena being explained, they nevertheless suggest that religious affiliations do not strictly divide the way that people explain the world.

Next, we tested whether between-group similarity varies across the domains of phenomena being explained. This tests our prediction that Christians and non-religious people use more similar concepts when describing natural phenomena than supernatural phenomena.

First, we tested whether between-group similarity varied more across domains than the similarity of non-religious people's explanations (Supplementary Table 3). We find evidence that between-group similarity varies less across natural and supernatural phenomena than the similarity of non-religious people's explanations (MCMCglmm: posterior mean = –14.71, CrI = –14.71 to 12.09, pMCMC < 0.001). When explaining the natural world, between-group similarity was lower than the similarity of non-religious people's explanations (Fig. 1). When explaining supernatural phenomena, between-group similarity was higher than the similarity of non-religious people's explanations (Fig. 1). Countervintuitive, this indicates that, when explaining supernatural phenomena, non-religious people proposed explanations that shared more concepts with the explanations of Christians than other non-religious people. This may be because non-religious people used a similar base set of concepts as Christians, along with a heterogeneous range of other concepts.

We also tested whether between-group similarity varied more across domains than the similarity of Christian's explanations (Supplementary Table 3). We found that between-group similarity varied more across natural and supernatural phenomena than the similarity of Christian's explanations (MCMCglmm: posterior mean = –28.75, CrI = –34.24 to –23.40, pMCMC < 0.001). When explaining supernatural phenomena, between-group similarity was close to the similarity of Christian's explanations (Fig. 1). This indicates that the concepts used by Christians to explain natural phenomena were largely overlapping with those used by non-religious people. Non-religious people proposed explanations that were based on a relatively narrow base of words, often involving science-based concepts (Supplementary Table 11). While Christians often refer to the same base concepts as non-religious individuals, some also use an expanded range of supernatural concepts. For example, when asked to explain why Earth's atmosphere contains oxygen and blocks UV radiation, a Christian participant responded “God created the earth as a home for humanity, the laws of nature in place keep the
balance.” This illustrates how Christian worldviews can synthesize science-based and religious-based concepts (Evans & Lane, 2011; Legare & Visala, 2011; Watts, 1997).

### 3.3. Testing the validity of Jaccard similarity

Inspection of our data showed that overlap in concepts did not always strictly indicate agreement between participants. For example, when explaining supernatural phenomena, the explanations proposed by non-religious participants sometimes contained negations, qualification, or implied a lack of endorsement. When asked to explain the contents of the Bible, a non-religious participant wrote “The Bible’s point is to guide people to a good life, no matter how or by whom it was written.” This highlights one of the limitations of using Jaccard similarity: while this measure of similarity gets at the broad conceptual overlap in explanations, it does not always capture the subtleties expressed in language.

To check the validity of our methodological approach, we performed a follow-up study testing whether Jaccard similarity reliably corresponds to human coder’s perception of similarity (Supplementary Study B). In the follow-up study, we had 100 additional participants rate the similarity of a subset of explanations proposed in the main study. Participants were presented two explanations for the same phenomena and asked to rate the similarity of these explanations on a Likert scale of 0 (Very different) to 10 (Very similar). We then tested whether the human coder’s ratings of similarity predicted our automated Jaccard similarity measure. Our results show that human similarity ratings significantly predicted Jaccard Similarity (MCMCglmm: posterior mean = −8.70, CRf = −11.70 to −5.67, pMCMC < 0.001), indicating that Jaccard similarity generally corresponds to how people perceive the similarity of different explanations.

### Table 1

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Posterior distribution mean</th>
<th>Lower 95% CrI</th>
<th>Upper 95% CrI</th>
<th>Effective sample size</th>
<th>pMCMC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Religion comparison</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Different religions</td>
<td>18.06</td>
<td>5.70</td>
<td>32.39</td>
<td>1000.00</td>
<td>0.006</td>
</tr>
<tr>
<td>Both Christian</td>
<td>24.18</td>
<td>1.13</td>
<td>53.24</td>
<td>1199.13</td>
<td>0.052</td>
</tr>
<tr>
<td>Domain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supernatural</td>
<td>48.01</td>
<td>9.77</td>
<td>87.37</td>
<td>1000.00</td>
<td>0.014</td>
</tr>
<tr>
<td>Age difference</td>
<td>0.22</td>
<td>0.06</td>
<td>0.37</td>
<td>1000.00</td>
<td>0.008</td>
</tr>
<tr>
<td>Gender difference</td>
<td>3.27</td>
<td>1.40</td>
<td>5.55</td>
<td>884.80</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Education difference</td>
<td>0.65</td>
<td>−0.70</td>
<td>2.11</td>
<td>854.53</td>
<td>0.356</td>
</tr>
<tr>
<td>Religion comparison: subject domain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Different religions: supernatural</td>
<td>−32.30</td>
<td>−36.50</td>
<td>−27.55</td>
<td>1000.00</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Both Christian: supernatural</td>
<td>−60.91</td>
<td>−66.82</td>
<td>−54.99</td>
<td>1000.00</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

Fig. 1. Across all parts of the image, blue represents Christian participants, red indicates non-religious participants, and purple represents the relationship between Christian and non-religious participants. The line graph in part A represents the mean similarity of participant explanations by domain of phenomena. Error bars represent 95% confidence intervals. The networks illustrated in Parts B and C represent the mean similarity of participants’ explanations for natural (Part B) and supernatural phenomena (Part C). The networks represent the strongest 300 links between participants, and participants not connected by any edges were removed from the network. The networks show only the participants that share the most similar explanations with others, and participants that propose more similar explanations to one another will tend to be closer together. Despite non-religious individuals proposing more similar explanations for natural phenomena, and Christians proposing more similar explanations for supernatural phenomena, there is substantial overlap between the concepts used by Christian and non-religious people across both networks. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
3.4. Additional analyses

We performed additional analyses to test whether participants self-identified affiliation of Christian and non-religious correspond to differences in commitments to supernatural beliefs. All participants in the main study completed the Revised Paranormal Belief Scale, and we used the Traditional Religious Belief dimension as a measure of Christian belief (Tobacyk, 2004). We found that Christians have greater commitment to Traditional Religious Beliefs (Mdn = 5.75) than non-religious participants (Mdn = 1.25), W = 10,822, p < .001 (Supplementary Table 15). This indicates that there are clear differences in the commitment of our non-religious and Christian participants to traditional religious beliefs (Supplementary Fig. 5).

In the Supplementary Methods section, we also report the results of further analyses testing how the sentiment of phenomena and the specific sub-domain of phenomena predict the similarity of participants’ explanations (Supplementary Tables 4–9).

4. Discussion

Our findings challenge the popular claim that religious system homogenize people's worldviews in general (Dawkins, 2006; Dennett, 2006). Instead, we find that the homogeneity of Christian and non-religious people's explanations depends on the kind of phenomena being explained. Christians proposed more homogenous explanations than non-religious people for supernatural phenomena, but not natural phenomena. When explaining the natural world, Christian and non-religious people primarily drew on science-based concepts, with some Christians supplementing these concepts with religious-based concepts. When explaining supernatural phenomena, Christians drew on a shared set of religious-based concepts, but non-religious people lacked a common conceptual framework and showed relatively little consensus in their explanations. This suggests that Christianity primarily provides a common conceptual framework for supernatural phenomena and that non-religious people have a diverse range of perspectives on the supernatural.

We also found substantial overlap in the concepts used by Christian and non-religious people. When explaining supernatural phenomena, non-religious people proposed explanations that, on average, shared more concepts with Christians than they did with other non-religious people. When explaining the natural world, the explanations proposed by Christians shared a similar number of concepts to the explanations of other Christians as they did to the explanations of non-religious people. Consistent with the predictions of synthesis-based accounts, this suggests that the primary difference between religious and secular world-views is in the scope of concepts drawn upon, rather than the core concepts (Legare & Visala, 2011; Watson-Jones, Busch, & Legare, 2015). This challenges the idea that religious and secular worldviews are necessarily competing and demonstrates how text analytics can efficiently quantify the structure and diversity of group ideologies.

Acknowledgments

This research was funded by the Templeton Religion Trust (TRT0153), who had no input to the study. We thank Cassie Sprules for assistance coding participant explanations, as well as the Editor and two anonymous reviewers for their constructive feedback

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2020.104290.

References

Nancekivell, S. E., & Friedman, O. (2017). She bought the unicorn from the pet store: Six-to-seven-year-olds are strongly inclined to generate natural explanations. Developmental Psychology, 53(6), 1079.