Exploring the Relationship Between Religion and Trust in Science

Enzo Moraes Mescall, Charlotte Coudert, Martin Olarte

2022-12-18

Abstract

The relationship between religiosity and acceptance of science is a topic of ongoing debate and investigation. Using data from the General Social Survey, this paper aims to examine the effect of religiosity on trust in science among adults in the United States. An ordinal logistic regression model was trained to predict trust in science based on several predictors, including frequency of attendance to religious services, frequency of prayer, and how fundamentalist the respondent was during adolescence. Results suggest that there is a negative relationship between religiosity and trust in science, with individuals who pray more frequently, show more devotion to the Bible, and were fundamentalist in their adolescence exhibiting lower levels of trust in science. These findings have implications for the promotion of scientific literacy and understanding among religious populations.

Introduction

In April 2020, President of the American Medical Association Dr. Patrice Harris said in a national address, "In these challenging times, I am addressing the nation to make an appeal for science in defeating this disease" [1]. The conflict between religion and science is a long standing issue in the U.S. but never quite so pertinent as during the COVID-19 pandemic. Belief in science and trust in national medical recommendations meant the difference between weeks of stringent lockdown or tentative freedom. It meant thousands of lives saved or lost.

Because of the contradictions between current standard teachings of science and traditional readings of the Bible and many other religions' stories of creation and human existence, it is no surprise that the news often pit the two in a 'war of faith and science.' Research on the topic is not sparse. There have been many correlational studies looking at specific relations like religion, religious attendance, and disbelief in evolution, the tendency for scientists to be less religious [2], and high religiosity and negative attitudes towards science [3].
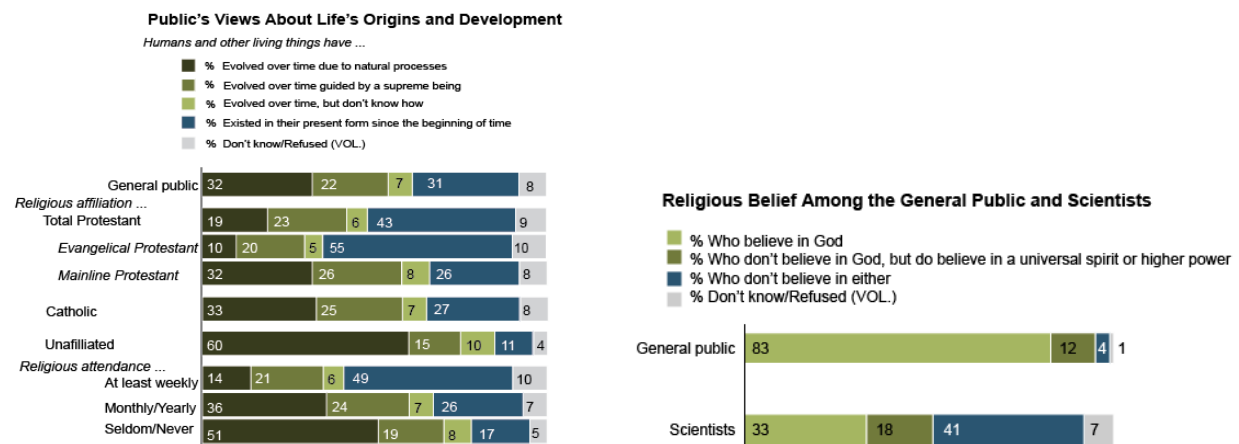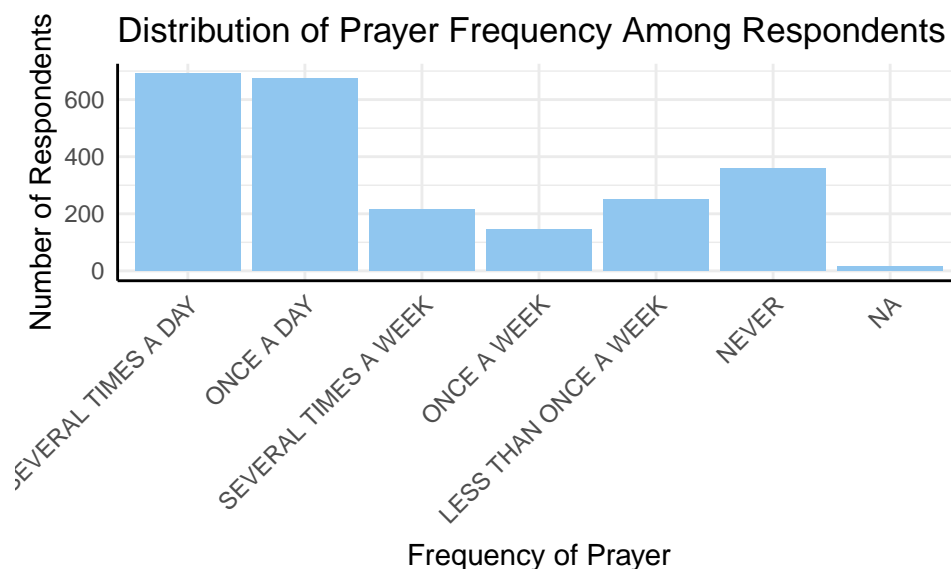
We aimed to add to the conversation with this project by exploring the relationship between religious beliefs and practices and trust in science. With the large scope of the data we used from the General Social Survey, we are able to investigate many facets of religiousness across multiple belief systems and look at their effects both individually and in interaction with each other. With this goal, we used logistic regression to see the individual effect of each factor and how they compare so the findings could be applicable on population trends, not just on an individual level.
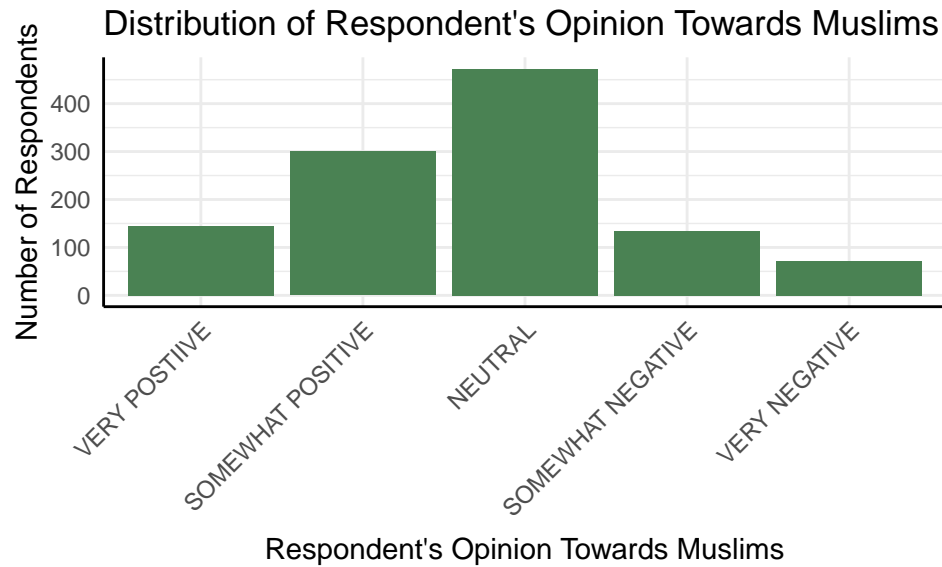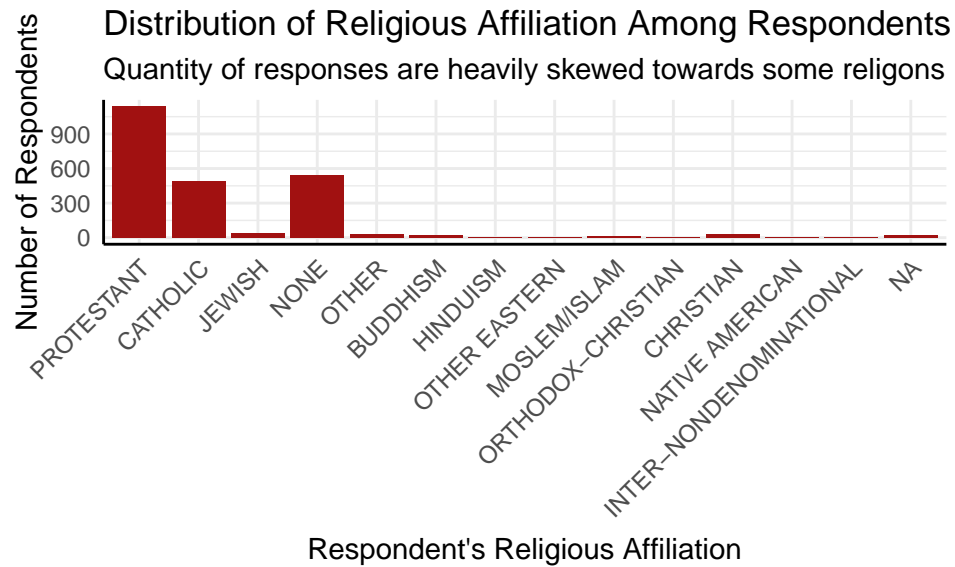
This topic is important to understand because with better understanding of how different demographics interact with science, we can approach public education more thoughtfully. Whether it is about diseases like COVID-19, global warming, public health recommendations, or emerging technologies, the language in which it is introduced can be crucial to its acceptance [4].
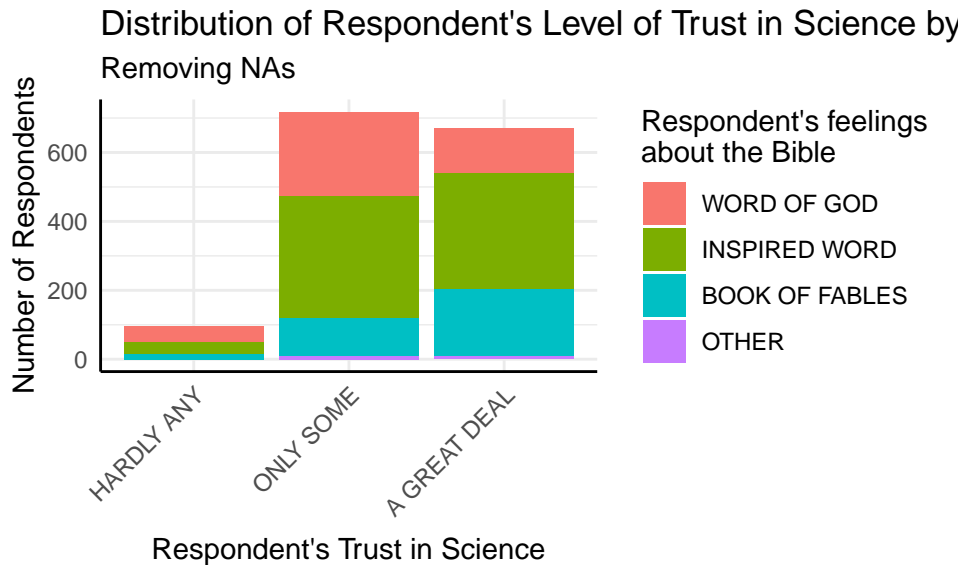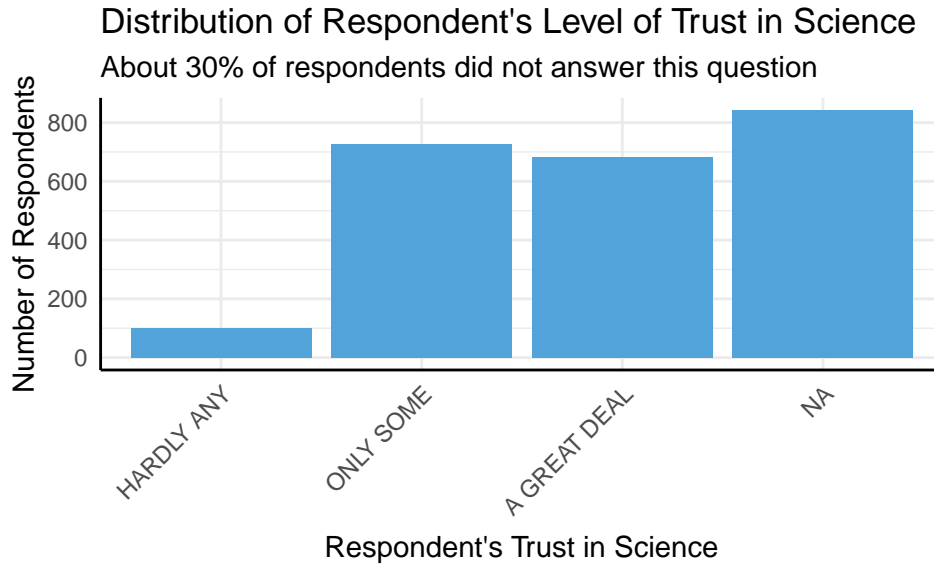
Data

The dataset in the project is the 2018 General Social Survey data, made available by NORC at the University of Chicago [5]. The General Social Survey (GSS) has been conducted most years since 1972 as a broad survey across the US containing "a standard core of demographic, behavioral, and attitudinal questions, plus topics of... civil liberties, crime and violence, intergroup tolerance, morality, national spending priorities, psychological well-being, social mobility, and stress and traumatic events" [5]. The breadth of this data allowed for us to consider many variables and focus on those that best answered our question both in interpretation and in statistical significance. For this survey, the data is collected in face-to-face interviews conducted by NORC and imputed by the interviewer for all the questions the respondent, "R" in the codebook, is able to answer. The value is "NA" for all values the respondent does not or cannot answer.

Exploratory Data Analysis

Distribution of Prayer Frequency Among Respondents

Distribution of Religious Affiliation Among Respondents

Quantity of responses are heavily skewed towards some religons

Respondent's Religious Affiliation

Distribution of Respondent's Opinion Towards Muslims

Respondent's Opinion Towards Muslims

Distribution of Respondent's Level of Trust in Science
About 30% of respondents did not answer this question

Number of Respondents

Respondent's Trust in Science

Distribution of Respondent's Level of Trust in Science by
Removing NAs

Number of Respondents

Respondent's feelings about the Bible

- WORD OF GOD
- INSPIRED WORD
- BOOK OF FABLES
- OTHER

Respondent's Trust in Science

Data Cleaning

First, we created a codebook for all the variables in the data by mapping the predictor names to their description. Then, to select relevant predictors we parsed the descriptions, looking for matches to religious key words and phrases like "relig", "god", " pray", etc. using the grep() function in R. Our resulting dataset (see Appendix 2) held 82 potential religious predictors and one response (CONSCI) which measured the respondent's confidence in the scientific community. Next, we filtered this dataset for predictors with less than 10% NA values since the nature of the survey led many predictors to have > 50% values. For example, certain questions like COJew asked "What sort of Jew is R's partner currently", which would only be answered by respondents who currently have a partner that is Jewish (in this case 99.91% of respondents had NA values). The new filtered dataset was composed of 2,348 observations with 12 predictors and 1 response. When fitting the first ordered logit model, we realized that all predictors except RELIGION (R's religious preference) were ordinal and therefore interpretable. Thus, to continue considering RELIGION we decided to divide the variable into individual binary predictors for each of the available options of the question.

However, when fitting an ordered logit model on the augmented dataset, the design matrix for the model was rank deficient. This can happen if the model includes variables that are perfectly correlated with one another, or if the model includes variables that are linearly dependent on a constant or an intercept term. In this case, there was a probable perfect correlation between one of the specific religions and another predictor with very few responses. A rank deficiency in the design matrix could lead to problems when fitting an ordered logit model, as it can result in an ill-posed optimization problem that does not have a unique solution which can cause the model fitting procedure to fail or produce unstable or unreliable estimates. Thus, we decided to remove all religions that would cause such a rank deficiency, and were left with new predictors for respondents who answered their preference was Christian, Catholic, Protestant, or None.

Finally, the last step in data cleaning was for the third iteration of the model. We releveled the response variable CONSCI to have more interpretable results, changing the baseline level to "Hardly any confidence in the scientific community" (CONSCI = 1), then "Only some confidence in the scientific community" (CONSCI = 2), and finally "A great deal of confidence in the scientific community" (CONSCI = 3). This made it possible to interpret coefficients in terms of the multiplicative effect on the odds of being more confident in the scientific community rather than less confident, which would be more convoluted.

Modeling

An ordinal logistic regression model is appropriate for this study because the response variable, (CONSCI) confidence in the scientific community, is an ordinal variable. An ordinal variable is a categorical variable that can be ordered or ranked, which is the case for CONSCI as seen above in the releveling of the variable.

In contrast, we avoided a multinomial logistic regression model since they are better suited to predict the probability of an individual falling into each category of a response variable that has three or more unordered categories. It is particularly useful when the categories of the response variable are unordered and there are a large number of categories, but as stated the ordinal logistic regression model allows us to account for the ordinal nature of the response variable and model the probability of an individual falling into each category based on their values of the predictors.

We also included confidence intervals for the parameter estimates. These can be obtained either by profiling the likelihood function or by using the standard errors and assuming a normal distribution. Note that profiled confidence intervals are not symmetric (although they are usually close to symmetric).

```
## Call:
## polr(formula = as.factor(CONSCI) ~ ., data = the_df, Hess = TRUE)
##
## Coefficients:
##             Value Std. Error t value
## PRAY     -0.12809    0.04723 -2.7118
## ATTEND   -0.06074    0.03127 -1.9423
## RELIG16   0.01872    0.05430  0.3448
## RELPERSN  0.03917    0.08990  0.4357
## RELIG    -0.01660    0.05179 -0.3205
## RELACTIV -0.03539    0.03311 -1.0689
## GOD       0.06206    0.05809  1.0684
## BIBLE    -0.53648    0.10556 -5.0821
## RELITEN   0.07624    0.08218  0.9278
## FUND16   -0.22966    0.10763 -2.1338
## FUND     -0.08837    0.11759 -0.7516
## POSTLIFE -0.15110    0.16783 -0.9003
##
## Intercepts:
##     Value   Std. Error t value
## 1|2 -2.1251  0.5979     -3.5543
```

```
## 2|3  0.9337  0.5989     1.5590
##
## Residual Deviance: 2075.455
## AIC: 2103.455
## (1112 observations deleted due to missingness)


##                  Value Std..Error    t.value      p_value significance
## PRAY      -0.12808811 0.04723284 -2.7118443 6.691001e-03          ***
## ATTEND    -0.06073943 0.03127268 -1.9422519 5.210661e-02            *
## RELIG16    0.01872119 0.05429870  0.3447815 7.302586e-01
## RELPERSN   0.03916526 0.08990017  0.4356528 6.630887e-01
## RELIG     -0.01659847 0.05179093 -0.3204900 7.485969e-01
## RELACTIV  -0.03538994 0.03310827 -1.0689154 2.851078e-01
## GOD        0.06206247 0.05809114  1.0683638 2.853564e-01
## BIBLE     -0.53647870 0.10556272 -5.0820847 3.733148e-07          ***
## RELITEN    0.07624481 0.08217915  0.9277878 3.535177e-01
## FUND16    -0.22966492 0.10763116 -2.1338144 3.285797e-02           **
## FUND      -0.08837424 0.11758677 -0.7515662 4.523120e-01
## POSTLIFE  -0.15110273 0.16782752 -0.9003454 3.679365e-01
## 1|2       -2.12511194 0.59790685 -3.5542526 3.790550e-04          ***
## 2|3        0.93368640 0.59890560  1.5589876 1.189993e-01


## Call:
## polr(formula = as.factor(CONSCI) ~ ., data = new_df, Hess = TRUE)
##
## Coefficients:
##               Value Std. Error t value
## PRAY       -0.12289    0.04747 -2.5887
## ATTEND     -0.05912    0.03138 -1.8838
## RELIG16     0.02441    0.05567  0.4384
## RELPERSN    0.04057    0.09009  0.4504
## RELIG      -0.08137    0.19150 -0.4249
## RELACTIV   -0.04374    0.03362 -1.3012
## GOD         0.07183    0.05866  1.2245
## BIBLE      -0.53052    0.10564 -5.0218
## RELITEN     0.01750    0.12017  0.1456
## FUND16     -0.21740    0.10947 -1.9859
## FUND       -0.10948    0.12697 -0.8622
## POSTLIFE   -0.15591    0.16803 -0.9278
## CHRISTIAN   0.35581    1.53744  0.2314
## CATHOLIC   -0.42247    0.60280 -0.7008
## NONE        0.03410    0.54563  0.0625
## PROTESTANT -0.30514    0.73071 -0.4176
##
## Intercepts:
##     Value   Std. Error t value
## 1|2 -2.5857  1.0741    -2.4073
## 2|3  0.4795  1.0732     0.4468
##
## Residual Deviance: 2073.141
## AIC: 2109.141
## (1112 observations deleted due to missingness)


## Waiting for profiling to be done...
```

```
##                OR     X2.5..    X97.5..
## PRAY    0.8843580 0.8056360 0.9705379
## ATTEND 0.9425972 0.8862224 1.0023069
## BIBLE   0.5883004 0.4775784 0.7227937
## FUND16 0.8046116 0.6489952 0.9971733


## Call:
## polr(formula = as.factor(CONSCI_re) ~ ., data = relevel_df, Hess = TRUE)
##
## Coefficients:
##               Value Std. Error  t value
## PRAY         0.12289    0.04747  2.58858
## ATTEND       0.05912    0.03138  1.88380
## RELIG16     -0.02441    0.05567 -0.43843
## RELPERSN    -0.04058    0.09009 -0.45041
## RELIG        0.08144    0.19151  0.42522
## RELACTIV     0.04374    0.03362  1.30102
## GOD         -0.07183    0.05866 -1.22457
## BIBLE        0.53052    0.10564  5.02179
## RELITEN     -0.01753    0.12017 -0.14590
## FUND16       0.21740    0.10947  1.98590
## FUND         0.10944    0.12697  0.86194
## POSTLIFE     0.15594    0.16803  0.92801
## CHRISTIAN   -0.35647    1.53753 -0.23185
## CATHOLIC     0.42247    0.60281  0.70084
## NONE        -0.03409    0.54564 -0.06248
## PROTESTANT   0.30519    0.73072  0.41766
##
## Intercepts:
##      Value    Std. Error t value
## 1|2 -0.4795   1.0733     -0.4468
## 2|3  2.5857   1.0742      2.4072
##
## Residual Deviance: 2073.141
## AIC: 2109.141
## (1112 observations deleted due to missingness)


## Waiting for profiling to be done...


##                OR     X2.5..   X97.5..
## PRAY    1.130758 1.0303565 1.241256
## ATTEND 1.060899 0.9976986 1.128385
## BIBLE   1.699812 1.3835206 2.093897
## FUND16 1.242841 1.0028347 1.540843
```

Inference

Our final model gave us information about 16 variables' effects on our predictor but only four were statistically significant. Holding all else constant, for every one unit decrease in frequency prayed, the odds of being more confident in science; here meaning being in a higher stratification of trust, e.g. very confident compared to somewhat confident; increases 13%. By the same definition, each unit increase in attendance to religious services, increases the odds of being more confident in science by 6%, each unit decrease in level of devotion

to the Bible increased the odds of being more confident in science by 70%, and each unit decrease in how fundamentalist the respondent was at age 16 increased odds of being more confident in science by 24%.

The confidence intervals for PRAY, BIBLE, and FUND16 do not include 0; ATTEND does. The estimates in the output are given in units of ordered logits, or ordered log odds. The coefficients from the model can be somewhat difficult to interpret because they are scaled in terms of logs. Thus, we can interpret logistic regression models by converting the coefficients into odds ratios (OR). To get the OR and confidence intervals, we just exponentiate the estimates and confidence intervals.

All of these conclusions but attendance support the literature we found on the topic and our initial hypothesis that higher levels of religiosity are correlated with lower confidence in science. The opposite effect of attendance is interestingly contrary to our expectations and other findings. It introduces a potential question for further research to see if this is specific to this data or if there is a general trend where religious attendance follows a different pattern than other signs of religiosity, potentially related to stronger correlations with community or free-time.

We note the smaller effects of frequency of prayer and inverse effect of attendance of religious services compared to those of devotion to the Bible and religious fundamentalism. While the former two imply attachment and devotion to religion, they do not mandate stringent belief of the teaching in the way the latter two do. This is a reasonable impact because the modern conflict between faith and science comes from their contradictions in much traditional religious teachings that fundamentalists and strict Bible followers would hold much more attachment to than other Americans, religious or not.

Among the variables not considered statistically significant, how fundamentalist the respondent was did not have statistical significance even though how fundamentalist they were at age 16 did. It is worth noting that statistical significance does not always equate to practical or real-world significance. In other words, even if a relationship is not statistically significant, it may still be meaningful or important in some way. It is always important to consider the size and strength of the relationship, as well as the context and potential implications of the finding, in addition to statistical significance. It is also worth considering the potential reasons for this finding. For example, it could be that religious fundamentalism at age 16 has a lasting impact on an individual's attitudes and beliefs, while religious fundamentalism at the current age may be less influential or more influenced by other factors. Alternatively, it could be that the sample size for religious fundamentalism at the current age was not large enough to detect a statistically significant relationship. Further research would be needed to better understand the implications of this finding.

Conclusion

In synthesis, in our analysis of the 2018 GSS data, collected across the United States using methods honed over almost 50 years of this survey, we found evidence of a relationship between multiple facets of religion and decreased confidence in science. The implications of this could mean a more precise version of what we introduced at the start of this paper. When scientists speak to each other, they speak of statistical significance, sample size, and reproducibility to convince others of their findings. However, when you leave those circles, those measures mean much less. We believe that science could be more effectively communicated to the general public if it was presented in a way most palatable to the targeted audience, even if they don't directly align with religious dogma. Based on these findings, which have been limited to data collected in 2018 and not across all general social surveys, it is implied that for more religious areas and groups, the scientific aspect could be played down in place of, in a case like COVID, promoting working together and protecting their community. More specifically, this tonal shift can be applied not just to where people are religious, because that is a majority of the Americans, but to where people are specifically fundamentalists or strict Bible followers. These findings show many people are willing to stray from their religion's teachings to the current science, as seen in the positive relationship between attendance and trust in science, but there are traceable, predictable, groups who are not willing to place faith in science, likely due to their religious beliefs.

In further research, it would be interesting to extend this model across the 50 years the survey has been conducted and see if the relations hold and if or how they have changed with time. Additionally, especially if

we can support our findings with the further data, an interesting follow-up study could test language choice in communicating the same information to groups of varying religiosity to see if our proposed impact of this paper's findings is viable.

Bibliography

- [1] "Covid-19: The Importance of Science in an Era of Distrust and Disinformation." American Medical Association, 8 Apr. 2020, https://www.ama-assn.org/press-center/speeches/covid-19-importance-science-era-distrust-and-disinformation.

- [2] Wormald, Benjamin. "Scientists and Belief." Pew Research Center's Religion & Public Life Project, Pew Research Center, 26 Apr. 2022, https://www.pewresearch.org/religion/2009/11/05/scientists-and-belief/.

- [3] McPhetres, Jonathon, and Miron Zuckerman. "Religiosity Predicts Negative Attitudes towards Science and Lower Levels of Science Literacy." PloS One, U.S. National Library of Medicine, 27 Nov. 2018, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6258506/.

- [4] Peters, Hans Peter. "Gap between Science and Media Revisited: Scientists as Public Communicators." PNAS, 12 Aug. 2013, https://www.pnas.org/doi/10.1073/pnas.1212745110.

- [5] "Get the Data ." General Social Survey, NORC at The University of Chicago, https://gss.norc.org/get-the-data/spss.