--- Metadata ---

Standard registration metadata questions

--- Part 1 ---

Study Information

[Question title]

Research Questions

[Description]

List each research question included in this study.

When specifying your research questions, it is good practice to use only two new concepts per research question. For example, split up your questions into a simple format: “Does X lead to Y?” and “Is the relationship between X and Y moderated by Z?”. By splitting up the research questions here, you can more easily describe the statistical test for each research question later.

[Example]

**Example:** RQ1 = Are more religious people more prosocial than less religious people? RQ2 = Does the relationship between religiosity and prosociality differ for people with different religious affiliations?

[Question title]

Hypotheses

[Description]

For each of the research questions listed in the previous section, provide one or more specific and testable hypothesis. Please make clear whether the hypotheses are directional (e.g., A > B) or non-directional (e.g., A ≠ B). If directional, state the direction. You may also provide a rationale for each hypothesis.

[Example]

We expect that religiosity is associated with prosocial behavior in our sample as well. To assess this prediction, we will use the following hypotheses:H0(1) = In men and women who graduated from Wisconsin high schools in 1957, there is no association between religiosity and prosociality H1(1) = In men and women who graduated from Wisconsin high schools in 1957, there is an association between religiosity and prosociality.

It is possible that these mechanisms are stronger for some religions than for others. and prosocial behavior differs for the different religions in our sample. To answer this question, we will test the following hypotheses: H0(2) = The relationship between religiosity and prosociality in men and women who graduated from Wisconsin high schools in 1957 does not differ for people with different religious affiliations (i.e., Roman Catholics, Lutheran, Methodists, Presbyterians, United Church of Christ, and Protestants) H1(2) = The relationship between religiosity and prosociality in men and women who graduated from Wisconsin high schools in 1957 differs for people with different religious affiliations (i.e., Roman Catholics, Lutheran, Methodists, Presbyterians, United Church of Christ, and Protestants).

--- Part 2 ---

Data description

[Question title]

Datasets Used

[Description]

Name and briefly describe the dataset(s), and if applicable, the subsets of the data you plan to use.

Useful information to include here is the type of data (e.g., cross-sectional or longitudinal), the general content of the questions, and some details about the respondents. In the case of longitudinal data, information about the survey’s waves is useful as well. Mention the most relevant information so that readers do not have to search for the information themselves.

[Example]

To answer our research questions we will use a dataset from the Wisconsin Longitudinal Study (WLS; Herd, Carr, & Roan, 2014). The WLS provides long-term data on a random sample of all the men and women who graduated from Wisconsin high schools in 1957. The WLS involves twelve waves of data. Six waves were collected from the original participants or their parents (1957, 1964, 1975, 1992, 2004, and 2011), four were collected from a selected sibling (1977, 1994, 2005, and 2011), one from the spouse of the original participant (2004), and one from the spouse of the selected sibling (2006). The questions vary across waves and are related to domains as diverse as socio-economic background, physical and mental health, and psychological makeup. We will use the subset consisting of the 1957 graduates who completed the follow-up 2003-2005 wave of the WLS dataset because it includes specific modules on religiosity and volunteering.

[Question title]

Data availability

[Description]

Specify the degree to which the datasets are open or publicly available.

[Responses]

[] The dataset is publicly available.

[] The dataset is available through protected access.

[] The dataset is not publicly available.

[Question Title]

Data Access

[Description]

If there are any restrictions to accessing the dataset, please describe this here.

[Example]

The dataset was made available through private access for the purposes of this study.

[Question Title]

Data Identifiers

[Description]

Please provide a URL, DOI, or other persistent, unique identifier of the dataset.

[Question Title]

Access Date

[Description]

Specify the download or data access date. If the data were accessed multiple times by different team members, specify the download date for that data that will be used in the statistical analysis.

[Example]

Downloaded 12 February 2019

[Question title]

Data Collection Procedures

[Description]

If the data collection procedure is well documented, provide a link to that information. If the data collection procedure is not well documented, describe, to the best of your ability, how data were collected. Describe the representativeness of the sample and any possible biases stemming from the data collection.

[Example]

The WLS data was collected by the University of Wisconsin Survey Center for use by the research community. The origins of the WLS can be traced back to a state-sponsored questionnaire administered during the spring of 1957 at all Wisconsin high school to students in their final year. Therefore, the dataset constitutes a specific sample not necessarily representative of the United States as a whole. Most panel members were born in 1939, and the sample is broadly representative of white, non-Hispanic American men and women who completed at least a high school education. A flowchart for the data collection can be found here: <https://www.ssc.wisc.edu/wlsresearch/about/flowchart/cor459d7.pdf>

[Question title]

Codebook

[Description]

Some studies offer codebooks to describe their data. If such a codebook is publicly available, link, cite, or upload the document. If not, provide other available documentation. Also provide guidance on what parts of the codebook or other documentation are most relevant.

[Example]

The codebook for the dataset we use can be found here: <https://www.ssc.wisc.edu/wlsresearch/documentation/waves/?wave=grad2k>. We will mainly use questions from the mail survey about religion and spirituality, and the phone survey on volunteering, but will also use some questions from other modules.

--- Part 3 ---

Variables

**[Question Title]**

Manipulated Variables

[Description]

If you are going to use any manipulated variables from the study variables, identify them here. Describe the variables and the levels or treatment arms of each variable. Note that this is not applicable for observational studies and meta-analyses. If you are collapsing groups across variables this should be explicitly stated, including the relevant formula. If your further analysis is contingent on a manipulation check, describe your decisions rules here.

[Example]

Not applicable

[Question Title]

Measured Variables

[Description]

Describe both outcome measures as well as predictors and covariates and label them accordingly. If you are using a scale or an index, state the construct the scale/index represents, which items the scale/index will consist of, and how these items will be aggregated. When the aggregation is based on exploratory factor analysis (EFA) or confirmatory factor analysis (CFA), also specify the relevant details (EFA: rotation, how the number of factors will be determined, how best fit will be selected, CFA: how loadings will be specified, how fit will be assessed, which residuals variance terms will be correlated). If you are using any categorical variables, state how you will code them in the statistical analyses.

[Example]

The following example is shortened. Please see <https://psyarxiv.com/hvfmr/> for complete details.

Religiosity (IV): Religiosity is measured using a newly created scale with a subset of items from the Religion and Spirituality module of the 2004 mail survey. The scale includes general questions about how religious/spiritual the individual is and how important religion/spirituality is to them. The specific variables are as follows:

il001rer: How religious are you?

il002rer: How spiritual are you?

il003rer: How important is religion in your life?

il004rer: How important is spirituality in your life?

il005rer: How important was it, or would it have been if you had children, to send your children for religious or spiritual instruction?

il006rer: How closely do you identify with being a member of a religious group?

The levels of all of these variables are indicated by a Likert scale with the following options: (1) Not at all; (2) Not very; (3) Somewhat; (4) Very; (5) Extremely, as well as ‘System Missing’ (the participant did not provide an answer) and ‘Refused’ (the participant refused to answer the question). Variable il006rer additionally includes the option ‘Don’t know’ (the participant stated that they did not know how to answer the question). We will use the average score (after omitting non-numeric and ‘Don’t know’ responses) on the variables as a measure of religiosity.

Prosociality (DV): We will use three measures of prosociality that measure three aspects of engagement in other oriented activities. The prosociality variables come from the Volunteering module of the 2004 phone survey.

gv103re: Did the graduate do volunteer work in the last 12 months?

This dichotomous variable assesses whether or not the participant has engaged in any volunteering activities in the last 12 months. The levels of this variable are yes/no. Yes will be coded as ‘1’, no will be coded as ‘0’.

[Question title]

Unit of Analysis

[Description]

Which units of analysis (respondents, cases, etc.) will be included or excluded in your study? Taking these inclusion and exclusion criteria into account, indicate the expected sample size of the data you’ll be using for your statistical analyses. If you have a research question about a certain group you may need to exclude participants based on one or more characteristics. Be very specific when describing these characteristics so that readers will be able to redo your moves easily.

[Example]

Initially, the WLS consisted of 10,317 participants. As we are not interested in a specific group of Wisconsin people, we will not exclude any participants from our analyses. However, only 7,265 participants filled out the questions on prosociality and the number of siblings in the phone survey and only 6,845 filled out the religiosity items in the mail survey (Herd et al., 2014). This corresponds to a response rate of 73% and 69% respectively. Because we do not know whether the participants that did the mail survey also did the phone survey, our minimum expected sample size is 10,317 \* 0.73 \* 0.69 = 5,297.

[Question title]

Missing Data

[Description]

What do you know about missing data in the dataset (i.e., overall missingness rate, information about differential dropout)? How will you deal with incomplete or missing data? Provide descriptive information, if available, on the amount of missing data for each variable you will use in the statistical analyses. Based on this information, provide a new expected sample size.

[Example]

The WLS provides a documented set of missing codes. Reasons for missing data include: ‘System missing’ refers to the number of participants that did not or could not complete the questionnaire. ‘Partial interview’ refers to the number of participants that did not get that particular question because they were only partially interviewed. The rest of the codes include ‘inappropriate,’ ‘refused,’ and ‘could not code.’

[Question title]

Statistical Outliers

[Description]

How will you define what a statistical outlier is in your data and what will you do when you encounter them? If you plan to remove outliers, provide a new expected sample size. If you expect to remove many outliers or if you are unsure about your outlier handling strategy, it is good practice to preregister analyses including and excluding outliers. Note that this will be the definitive expected sample size for your study and you will use this number to do any power analyses.

[Example]

The dataset probably does not involve any invalid data since the dataset has been previously cleaned by the WLS data controllers and any clearly unreasonably low or high values have been removed from the dataset. However, to be sure we will create a box and whisker plot for all continuous variables (the dependent variables gv109re and gv111re, the covariate gk067ss, and the scale for religiosity) and remove any data point that appears to be more than 1.5 times the IQR away from the 25th and 75th percentile. Based on normally distributed data, we expect that 2.1% of the data points will be removed this way, leaving 1,358 out of 1,387 participants for the binary regression with gv103re as the outcome variable and 1,086 out of 1,109 participants, and 1,041 out of 1,063 participants for the linear regressions with gv109re and gv111re as the outcome variables, respectively.

[Question title]

Sampling Weights

[Description]

Are there sampling weights available with this dataset? If so, are you using them or are you using your own sampling weights? Sampling weights can be useful in secondary data analysis because the sample may not be entirely representative of the population you are interested in.

[Example]

The WLS dataset does not include sampling weights and we will not use our own sampling weights as we do not seek to make any claims that are generalizable to the national population.

--- Part 4 --- Knowledge of data

[Question title]

Prior Publication/Dissemination

[Description]

List the publications, working papers, and conference presentations you have worked on that are based on the dataset you will use. For each work, list the variables you analyzed, but limit yourself to variables that are relevant to the proposed analysis. If the dataset is longitudinal, also state which wave of the dataset you analyzed. Specify the previous works for each co-author separately.

Listing previous works based on the data also helps to prevent a common practice identified by the American Psychological Association (2019) as unethical: the so-called “least publishable unit” practice (also known as “salami-slicing”), in which researchers publish multiple papers on closely related variables from the same dataset.

[Example]

Both authors (PS and JC) have previously used the Graduates 2003-2005 wave to assess the link between Big Five personality traits and prosociality. The variables we used to measure the Big Five personality traits were ih001rei (extraversion), ih009rei (agreeableness), ih017rei (conscientiousness), ih025rei (neuroticism), and ih032rei (openness). The variables we used to measure prosociality were ih013rer (“To what extent do you agree that you see yourself as someone who is generally trusting?”), ih015rer (“To what extent do you agree that you see yourself as someone who is considerate to almost everyone?”), and ih016rer (“To what extent do you agree that you see yourself as someone who likes to cooperate with others?). We presented the results at the ARP conference in St. Louis in 2013 and we are currently finalizing a manuscript based on these results.

[Question title]

Prior Knowledge

[Description]

What prior knowledge do you have about the dataset that may be relevant for the proposed analysis? Your prior knowledge could stem from working with the data first-hand, from reading previously published research, or from codebooks. Provide prior knowledge for every author separately.

Indirect knowledge about the hypothesized association does not preclude a confirmatory analysis but should be transparently reported in this section. However, direct knowledge about the association between the variables in your hypothesis may indicate that you are unable to make unbiased analytic decisions to test this hypothesis.

[Example]

In a previous study (mentioned in Q17) we used three prosociality variables (ih013rer, ih015rer, and ih016rer) that may be related to the prosociality variables we use in this study. We found that ih013rer, ih015rer, and ih016rer are positively associated with agreeableness (ih009rec). Because previous research (on other datasets) shows a positive association between agreeableness and religiosity (Saroglou, 2002) there may be an effect of the agreeableness variable on our hypothesized association between religiosity and prosociality. To account for this effect we will include agreeableness in our analysis as a control variable. We did not find any associations between prosociality and the other Big Five variables.

--- Part 5 --- Analyses

[Question title]

Statistical Models

[Description]

For each hypothesis, describe the statistical model you will use to test the hypothesis. Include the type of model (e.g., ANOVA, multiple regression, SEM) and the specification of the model. Specify any interactions and post-hoc analyses and remember that any test not included here must be labeled as an exploratory test in the final paper.

[Example]

Our first hypothesis will be tested using three analyses since we use three variables to measure prosociality. For each, we will run a directional null hypothesis significance test to see whether a positive effect exists of religiosity on prosociality. For the first outcome (gv103re: Did the graduate do volunteer work in the last 12 months?) we will run a logistic regression with religiosity, the number of siblings, agreeableness, religious affiliation, and the interaction between religiosity and religious affiliation as predictors.

For the second and third outcomes (gv109re: Number of graduate’s other volunteer activities in the past 12 months; gv111re: How many hours did the graduate volunteer during a typical month in the last 12 months?) we will run two separate linear regressions with religiosity, the number of siblings, agreeableness, religious affiliation, and the interaction between religiosity and religious affiliation as predictors.

If the regression coefficient of the interaction between religiosity and religious affiliation is statistically significant from zero in any of the regressions, we will also carry out subgroup analyses for those particular regressions in line with our second hypothesis. This subgroup analysis involves running the regressions separately for each of the following groups: Roman Catholics, Lutherans, Methodists, Presbyterians, and followers of the United Church of Christ.

The code we will use for all these analyses can be found at <https://osf.io/e3htr>.

[Question title]

Effect Size

[Description]

If applicable, specify a predicted effect size or a minimum effect size of interest for all the effects tested in your statistical analyses.

[Example]

For the logistic regression with ‘Did the graduate do volunteer work in the last 12 months?’ as the outcome variable, our minimum effect size of interest is an odds of 1.05. This means that a one-unit increase on the religiosity scale would be associated with a 1.05 factor change in odds of having done volunteering work in the last 12 months versus not having done so.

For the linear regressions with ‘The number of graduate’s volunteer activities in the last 12 months”, and “How many hours did the graduate volunteer during a typical month in the last 12 months?’ as the outcome variables, the minimum regression coefficients of interest of the religiosity variables are 0.05 and 0.5, respectively. This means that a one-unit increase in the religiosity scale would be associated with 0.05 extra volunteering activities in the last 12 months and with 0.5 more hours of volunteering work in the last 12 months.

All of these smallest effect sizes of interest are based on our own intuition.

[Question title]

Statistical Power

[Description]

Present the statistical power available to detect the predicted effect size or the smallest effect size of interest. Use the sample size after updating for missing data and outliers.

[Example]

The sample size after updating for missing data and outliers is 1,358 for the logistic regression with gv103re as the outcome variable, and 1,086 and 1,041 for the linear regressions with gv109re and gv111re as the outcome variables, respectively. For all three analyses this corresponds to a statistical power of approximately 1.00 when assuming our minimum effect sizes of interest. For the linear regressions we additionally assumed the variance explained by the predictor to be 0.2 and the residual variance to be 1.0. For the logistic regression we assumed intercept of -1.56 corresponding to a situation where half of the participants have done volunteer work in the last year (see the R-code for the full power analysis at <https://osf.io/f96rn>).

[Question title]

Inference Criteria

[Description]

What criteria will you use to make inferences? Describe the information you will use (e.g. specify the p-values, effect sizes, confidence intervals, Bayes factors, specific model fit indices), as well as cut-off criteria, where appropriate. Will you be using one-or two-tailed tests for each of your analyses? If you are comparing multiple conditions or testing multiple hypotheses, will you account for this, and if so, how?

[Example]

We will make inferences about the association between religiosity and prosociality based on the p-values and the size of the regression coefficients of the religiosity variable in the three main regressions. We will conclude that a regression analysis supports our hypothesis if both the p-value is smaller than .01 and the regression coefficient is larger than our minimum effect size of interest.

With regard to the second research question, we will conclude that religious groups differ in the association between religiosity and prosociality if we find a statistically significant coefficient using a two-tailed test of religious affiliation in at least one of the regressions.

We will not use any inference criteria about the follow-up analyses per religious group because we are not going to draw any conclusions based on these analyses.

[Question title]

Assumption Violation/Model Non-Convergence

[Description]

What will you do should your data violate assumptions, your model not converge, or some other analytic problem arises?

**[Example:** When the distribution of the number of volunteering hours (gv111re) is significantly non-normal according to the Kolmogorov-Smirnov test and/or (b) the linearity assumption is violated (i.e., the points are asymmetrically distributed around the diagonal line when plotting observed versus the predicted values), we will log-transform the variable.

[Question title]

Reliability and Robustness Testing

[Description]

Provide a series of decisions or tests about evaluating the strength, reliability, or robustness of your finding. This may include within-study replication attempts, additional covariates, cross-validation, applying weights, selectively applying constraints in an SEM context (e.g., comparing model fit statistics), overfitting adjustment techniques used, or some other simulation/sampling/bootstrapping method.

[Example]

To assess the sensitivity of our results to our selection criterion for outliers, we will run an additional analysis without removing any outliers.

[Question title]

Exploratory Analysis

[Description]

If you plan to explore your dataset to look for unexpected differences or relationships, describe those tests here. If reported, add them to the final paper under a heading that clearly differentiates this exploratory part of your study from the confirmatory part.

[Example]

As an exploratory analysis, we will test the relationship between scores on the religiosity scale and prosociality after adjusting for a variety of social, educational, and cognitive covariates that are available in the dataset. We have no specific hypotheses about which covariates will attenuate the religiosity-prosociality relation most substantially, but we will use this exploratory analysis to generate hypotheses to test in other, independent datasets.