**Preregistration Form: Analysis Plan**

**1.  Title**

Many Analysts Replication Project

**2.  Analysis Team**

|  |  |  |
| --- | --- | --- |
| **Surname, Name** | **Affiliation** | **Email-address** |
| Olsson Collentine, Anton | Tilburg University | anton@olssoncollentine.com |
| Gielens, Erwin | Tilburg University | erwin.gielens@gmail.com |
|  |  |  |

**3.  Research Question**

a.    Do religious people have higher well-being?

b.    Does the relationship between religiosity and well-being depend on perceived cultural norms of religion?

**4.  Hypotheses**

*List specific, concise, and testable hypotheses. Please state if the hypotheses are directional or non-directional. If directional, state the direction. A predicted effect is also appropriate here. If a specific interaction or moderation is important to your research, you can list that as a separate hypothesis.*

*Example*

*If taste affects preference, then mean preference indices will be higher with higher concentrations of sugar.*

|  |
| --- |
| 1. *People with higher religiosity have higher well-being on average* 2. *If the importance of religion in society is perceived to be lower, the effect of religiosity on well-being is more positive.* |

**5.  Variables**

**a.    Dependent variable(s)**

*State which key dependent variable(s) you will use in your analysis. Name the specific column names of these variables as stated in the data documentation.*

|  |
| --- |
| *We will create 4 scales of well-being, as defined in the data documentation. These 4 scales together will contribute to the latent construct ‘well-being’. The scales are defined as:*  ***General well-being:***  *14. wb\_gen\_1 quality of life general*  *15. wb\_gen\_2 satisfaction with health*  ***Physical well-being***  *16. wb\_phys\_1 pain [R]*  *17. wb\_phys\_2 medical dependence [R]*  *18. wb\_phys\_3 energy*  *19. wb\_phys\_4 mobility*  *20. wb\_phys\_5 sleep*  *21. wb\_phys\_6 activities*  *22. wb\_phys\_7 work ability*  *23. wb\_psych\_1 enjoying life*  ***Psychological well-being***  *23. wb\_psych\_1 enjoying life*  *24. wb\_psych\_2 meaningfulness*  *25. wb\_psych\_3 concentration*  *26. wb\_psych\_4 satisfaction physical appearance*  *27. wb\_psych\_5 self-esteem*  *28. wb\_psych\_6 negative affect [R]*  ***Social well-being***  *29. wb\_soc\_1 personal relations*  *30. wb\_soc\_2 social support*  *Note: we exclude the item*  *31. wb\_soc\_3 sexual satisfaction*  *From the social well-being scale because of its many missing values*  *These are then combined through SEM into the construct ‘well-being’.* |

**b.    Predictor variable(s)**

*State which predictor variable (including moderators and covariates) you will use in your analysis. Name the specific column names of these variables as stated in the data documentation.*

|  |
| --- |
| *As our primary predictor, we will create a latent construct indicated by all religiosity items:*  ***Religiosity***  *3. rel\_1 frequency of service attendance (1-7) [R]*  *4. rel\_2 frequency of prayer (1-8) [R]*  *5. rel\_3 self-identification (1= religious, 2= not religious, 3=atheist) [R]*  *6. rel\_4 belong to denomination (1=yes, 2=no) [R]*  *7. rel\_5 belief in God/Gods (1-7)*  *8. rel\_6 belief in afterlife (1-7)*  *9. rel\_7 spirituality (1-7)*  *10. rel\_8 importance of religious lifestyle (1-5)*  *11. rel\_9 importance of belief in God/Gods (1-5)*  ***Moderator***  *Perceived cultural norms of religiosity, a latent construct indicated by two items:*  *12. cnorm\_1 importance of religious lifestyle for average person in country (1-5)*  *13. cnorm\_2 importance of belief in God/Gods for average person in country (1-5)*  ***Covariates***  *42. gdp gross domestic product per capita (in US$, data from 2017; country-level)*  *44. sample\_type type of sample (general public, students, online panel, mixed)*  *C. We will create dummies for the countries and use as a covariate*  *36. age age in numbers*  *37. gender (1=man, 2=woman, 3=other)*  *38. ses subjective socioeconomic status as ladder (1 = bottom, 10 = top)*  *39. education education in levels (largely equivalent across samples; 1-7, 1 indicating*  *the lowest level of education, 7 indicting the highest level of education)* |

**c.    Indices**

*If applicable, please define how measures will be combined into an index (or even a mean) and what measures will be used. Include either a formula or a precise description of the method. If you are using a more complicated statistical method to combine measures (e.g. a factor analysis), please note that here but describe the exact method in the analysis plan section.*

*Example*

*We will take the mean of the two questions above to create a single measure of ‘brownie enjoyment.’*

|  |
| --- |
| *We will create latent constructs of Religiosity, perceived cultural norm of religiosity, general well-being, physical well-being, psychological well-being, social well-being, and ‘well-being’ using Structural Equation Modelling.* |

**6.  Analysis Plan**

**a.    Statistical models**

*What statistical model will you use to test each hypothesis? Please include the type of model (e.g. ANOVA, RMANOVA, MANOVA, multiple regression, SEM, etc) and the specification of the model. This includes each variable that will be included, all interactions, subgroup analyses, pairwise or complex contrasts, and any follow-up tests from omnibus tests. Provide enough detail so that another person could run the same analysis with the information provided. Remember that in your final article any test not included here must be noted as exploratory and that you must report the results of all tests.*

***Note:*** *This is perhaps the most important and most complicated question within the preregistration. Ask yourself: is enough detail provided to run the same analysis again with the information provided by the user? Be aware for instances where the statistical models appear specific, but actually leave openings for the precise test.*

*Example*

*We will use a  2 X 3 repeated measures ANOVA (RMANOVA) with the mean preference indices as the outcome variable and the factors “sweetness” and “color” within subjects to analyze our results.*

|  |
| --- |
| *We will use a SEM model with country as dummy variables, and well-being as the outcome variable. This model consists, broadly of:*  *i) one latent factor indicated by items rel\_1:rel\_9*  *ii) one latent factor for each of general, physical, psychological, and social well-being indicated by their respective items*  *iii) These will in turn be indicators of the latent variable 'well-being'*  *iv) a latent factor for perceived cultural norms, to moderate the relationship between religiosity and well-being*  *Below we present the code for fitting our model in R using the package lavaan.*  *To test the main effect (hypothesis 1), we fit the following model in lavaan:*  *m\_sem\_direct <- '*  *rel =~ rel\_1 + rel\_2 + rel\_3 + rel\_4 + rel\_5 + rel\_6 + rel\_7 + rel\_8 + rel\_9*  *wb\_gen =~ gen\_1 + gen\_2*  *wb\_phys =~ phys\_1 + phys\_2 + phys\_3 + phys\_4 + phys\_5 + phys\_6 + phys\_7*  *wb\_psych =~ psych\_1 + psych\_2 + psych\_3 + psych\_4 + psych\_5 + psych\_6*  *wb\_soc =~ soc\_1 + soc\_2*  *wb =~ wb\_gen + wb\_phys + wb\_psych + wb\_soc*    *wb ~ rel +*  *age + gender\_d1 + gender\_d2 +*  *education +*  *ses + gdp +*  *sample\_type\_d1 + sample\_type\_d2 + sample\_type\_d3 ++*  *country\_1 + country\_2 + country\_3 + country\_4 + country\_5 +*  *country\_6 + country\_7 + country\_8 + country\_9 + country\_10 +*  *country\_11 + country\_12 + country\_13 + country\_14 + country\_15 +*  *country\_16 + country\_17 + country\_18 + country\_19 + country\_20 +*  *country\_21 + country\_22 + country\_23*    *#fix factor variance for identification purposes*  *wb\_soc ~~ 1\*wb\_soc*  *wb\_gen ~~ 1\*wb\_gen*  *'*  *fit\_sem\_direct <- sem(m\_sem\_direct, data=dat\_sim)*  *summary(fit\_sem\_direct, fit.measures=T, rsquare=T, standardized = T)*  *To test the interaction effect we run the following code in lavaan:*  *#First fit cfa to compute the latent interaction variable*  *#Confirmatory Factor Analysis*  *m\_cfa <- 'rel =~ rel\_1 + rel\_2 + rel\_3 + rel\_4 + rel\_5 + rel\_6 + rel\_7 + rel\_8 + rel\_9*  *cnorm =~ cnorm\_1 + cnorm\_2*  *wb\_gen =~ gen\_1 + gen\_2*  *wb\_phys =~ phys\_1 + phys\_2 + phys\_3 + phys\_4 + phys\_5 + phys\_6 + phys\_7*  *wb\_psych =~ psych\_1 + psych\_2 + psych\_3 + psych\_4 + psych\_5 + psych\_6*  *wb\_soc =~ soc\_1 + soc\_2*  *wb =~ wb\_gen + wb\_phys + wb\_psych + wb\_soc*    *#fix factor variance for identification purposes*  *cnorm ~~ 1\*cnorm*  *wb\_soc ~~ 1\*wb\_soc*  *wb\_gen ~~ 1\*wb\_gen*  *'*  *fit\_cfa <- cfa(m\_cfa, data = dat\_sim, orthogonal=T)*  *summary(fit\_cfa, fit.measures=T, rsquare=T, standardized = T)*  *dat\_sim2 <- data.frame(dat\_sim, predict(fit\_cfa))*  *dat\_sim$rel\_cnorm <- dat\_sim2$rel \* dat\_sim2$cnorm #compute the latent interaction from the estimate latent variables*  *m\_sem\_int <- '*  *rel =~ rel\_1 + rel\_2 + rel\_3 + rel\_4 + rel\_5 + rel\_6 + rel\_7 + rel\_8 + rel\_9*  *wb\_gen =~ gen\_1 + gen\_2*  *wb\_phys =~ phys\_1 + phys\_2 + phys\_3 + phys\_4 + phys\_5 + phys\_6 + phys\_7*  *wb\_psych =~ psych\_1 + psych\_2 + psych\_3 + psych\_4 + psych\_5 + psych\_6*  *wb\_soc =~ soc\_1 + soc\_2*  *wb =~ wb\_gen + wb\_phys + wb\_psych + wb\_soc*    *cnorm =~ cnorm\_1 + cnorm\_2*    *wb ~ rel + cnorm + rel\_cnorm +*  *age + gender\_d1 + gender\_d2 +*  *education +*  *ses  + gdp +*  *sample\_type\_d1 + sample\_type\_d2 + sample\_type\_d3 +*  *country\_1 + country\_2 + country\_3 + country\_4 + country\_5 +*  *country\_6 + country\_7 + country\_8 + country\_9 + country\_10 +*  *country\_11 + country\_12 + country\_13 + country\_14 + country\_15 +*  *country\_16 + country\_17 + country\_18 + country\_19 + country\_20 +*  *country\_21 + country\_22 + country\_23*    *#fix factor variance for identification purposes*  *cnorm ~~ 1\*cnorm*  *wb\_soc ~~ 1\*wb\_soc*  *wb\_gen ~~ 1\*wb\_gen*  *'*  *fit\_sem\_int <- sem(m\_sem\_int, data=dat\_sim)*  *summary(fit\_sem\_int, fit.measures=T, rsquare=T, standardized = T)* |

**b.    Transformations**

*If you plan on transforming, centering, recoding the data, or requiring a coding scheme for categorical variables, please describe that process.*

*Example*

*The “Effect of sugar on brownie tastiness” does not require any additional transformations. However, if it were using a regression analysis and each level of sweet had been categorically described (e.g. not sweet, somewhat sweet, sweet, and very sweet), ‘sweet’ could be dummy coded with ‘not sweet’ as the reference category.  If any categorical predictors are included in a regression, indicate how those variables will be coded (e.g. dummy coding, summation coding, etc.) and what the reference category will be.*

|  |
| --- |
| *The variable ‘gender’ will be dummy-coded with 1 = man as the reference category.*  *37. gender (1=man, 2=woman, 3=other)*  *The variable ‘sample\_type’ will be dummy-coded with ‘general public’  as the reference category.*  *44. sample\_type (general public, students, online panel, mixed)*  *Country id will be recoded into a set of dummy variables. France will be the reference category.*  *We will not dummy-code any other variables.* |
|  |

**c.    Inference criteria**

*What criteria will you use to make inferences? Please describe the information you’ll use (e.g. specify the p-values, Bayes factors, specific model fit indices), as well as cut-off criterion, 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?*

*Example*

*We will use the standard p<.05 criteria for determining if the ANOVA and the post hoc test suggest that the results are significantly different from those expected if the null hypothesis were correct. The post-hoc Tukey-Kramer test adjusts for multiple comparisons.*

|  |
| --- |
| *Two effects of interest:*  *Direct effect of religiosity on wellbeing*  *Interaction effect of cultural norms on direct effect*  *Both effects are deemed significant if the two-sided p-values of the relevant regression parameters are <.05, indicating a significant deviation from a null effect.*    *In the first case the relevant regression parameter is the effect of ‘rel’ on ‘wb’ accounting for covariates as specified.*  *In the second case the relevant regression parameter is the effect of rel\_cnorm accounting for covariates.* |

**d.    Data exclusion**

*How will you determine which data points or samples if any to exclude from your analyses, for instance, based on the attention check or missing data? How will outliers be handled?*

*Example*

*We will verify that each subject answered each of the three tastiness indices. Outliers will be included in the analysis.*

|  |
| --- |
| *We will exclude any participants who failed the attention check item, i.e., did not score a 1 on the item below:*  *46. attention\_check attention check (1=passed)*  *In addition, participants with missing on any variable in our model (i.e,. age and SES) will be list-wise deleted. No other participants will be excluded.* |

**e.    Missing data**

*How will you deal with incomplete or missing data?*

***Note:*** *For the well-being and religiosity measures there are no missing data. See the data documentation for a more detailed overview of missing data for each variable.*

*Example*

*If a subject does not complete any of the three indices of tastiness, that subject will not be included in the analysis.*

|  |
| --- |
| *Participants with missing on any variable in our model (i.e.,  age and SES) will be list-wise deleted.* |

**7.  Other**

***If there is any additional information that you feel needs to be included in your preregistration, please enter it here. Literature cited, disclosures of any related work such as replications or work that uses the same data, or other helpful context would be appropriate here.***

|  |
| --- |
| Our analysis approach is based on the following preliminary assumptions:  a) that it makes sense to measure religiosity as a single factor construct (implied by the research question and data documentation)  b) that the religiosity items are all indicators of this construct (implied by the data documentation presenting all religiosity items without dividing them into subscales, as is done for well-being. Our read of the literature otherwise suggests that (at least) the subscales 'behavior', 'belief', and 'belonging' might be worth considering).  Hence, although we believe the research question is too broadly construed based on the theoretical background and would not formulate this research question ourselves, we will attempt to answer it based on the above assumptions. |