Examining the relationship between the big-5 personality facets and implicit racial attitudes

Details of the assignment

Ian Hussey

01 Januar, 2024

Contents

The data set was produced by a study which asked each participant to complete (a) between 2 and 3 of the big-5 personality subscales, (b) an Implicit Association Test assessing implicit racial evaluations between African-American and European-American people, an (c) demographics data including age and gender. Note that providing demographics data was optional, so missing demographics data does not constitute incomplete data.

Your assignment is to produce a well-organised and well-documented folder of data and code, including a script to process the data, a codebook for that processed data, and a script to analyse the data and report results. This code and data should maximise reproducibility so that it can run on someone else's machine simply by cloning your github repo and re-running the code.

Structure

- Ensure that your folder is called "personality and race-IAT study" and that it contains folders called "code", "data", and "communications" folders. "data" should contain "raw" and "processed" folders and their appropriate files.
- Ensure that your code folder contains files called processing.Rmd and analysis.Rmd. When you are finished, ensure that you knit these to also create processing.html and analysis.html.
- Your code should make use of consistent variable and object naming, eg using snake case (e.g., names_like_this) or whatever your preferred naming format is. Snake case is the default for tidyverse and I recommend it.
- Your code should make use of comments - it doesn't have to be exhaustive but provide some commentary on what you're doing, especially where code is more complex.
- Your code should make use of headings and levels of heading (i.e., using ###s) so that the file can be easily navigated.

Processing

- Extract age and gender from the raw demographics data.
- Reverse score the negatively worded items: the extroversion scale items 2, 5 and 7, conscientiousness items 2, 4 5 and 9, neuroticism items 2, 5, and 7, agreeableness 1, 3, 6, and 8, and openness items 7 and 9.
  - Include a sanity check that assesses whether these list of item reversals, and your implementation of them, is likely to be correct: For each subscale, create a correlation table among the items (after reversals) and check that all correlations are positive. IN general, negative correlations among items are one indication that information about item reversals, or their implementation, is not correct.
  - Check that the item level data does not violate the logical minimum and maximum scores (1 to 6). Create an exclusion variable and set participants with impossible data to "exclude".
  - Check that all participants have complete data on the BFI scales they completed. Create an exclusion variable and set participants with incomplete data to "exclude".
- Mean-score the subscales of the BFI scale. Each participant only got either 2 or 3 subscales.
  - Check that the mean scores do not violate the min and max possible score (i.e., first determine this min and max score), and revise your scoring code if it does.
- Score the trial-level IAT data using the Greenwald "D" score: Calculate a mean RT ("mean1") for blocks 3 and 6 (one score using trials from both blocks), a mean RT ("mean2") for blocks 4 and 7 (one score using trials from both blocks), and the SD of RTs in blocks 3, 4, 6 and 7 ("SD"). To calculate D: D = (mean2 - mean1)/SD. Blocks 1, 2, and 5 are practice blocks and must be discarded.
  - Include a sanity check: check that all D scores are in the range -2 to +2. If not, revise your implementation of the scoring code.
  - Create an exclusion variable and set participants with incomplete trial level IAT data to "exclude". Specifically, IAT should have 120 trials on the critical test blocks (i.e., blocks 3, 4, 6 and 7). Trials on the other (practice) blocks can be discarded.
  - Create an exclusion variable for IAT performance: set participants with >10% of the participants trials are < 300ms, or if their accuracy is < than 75%. Only use trials from the critical test blocks when computing these (i.e., blocks 3, 4, 6 and 7).
- Combine the demographics, BFI, and IAT data into one data frame. This data frame should be one-row-one-participant. Both the mean scored and item level BFI data should be present in the dataset.
- Create a master exclude variable from the individual exclude variables.
- Save the processed data to the data/processed/ folder as "data_processed.csv".
- Create a codebook for the processed data file that explains what each variable represents.

Analyses

- Create a table with the total number of participants before exclusions.
- Exclude participants using the master exclusion variable.
- Create a table with the total number of participants after exclusions.
- Create demographics tables that summarise the mean and SD age, and the frequency and percentage of men vs women in the sample. Report the proportion of participants for which each demographics variable was available.
- Calculate Cronbach's alpha for each of the BFI subscales.
- Create series of histograms of the distribution of scores on each of the BFI subscales and the IAT.
- Create a table with a correlation matrix of the Pearson's r correlations between the IAT and the BFI subscales rounded to two decimal places.

- Run a t-test to test the hypothesis that men and women differ on their scores on the IAT. Report and interpret the results using {report}.
- Run a series of regressions to test the hypotheses that each BFI subscale predicts IAT scores. Run separate regressions for each subscale rather than a multiple regression. Report and interpret the results using {report}.
- Create series of scatter plots of how BFI subscale scores predict IAT scores. Add regression lines. These plots therefore illustrate the same results as the regressions above. Make the plots presentable and publication quality- you may decide the details of this with one exception: Extreme scores on the IAT (defined as participants whose scores in the top and bottom 10% of IAT scores) should use a triangular shape rather than a dot. A single combined plot should be created, containing 5 component plots, with the personality facet being examined in each clearly labelled. This plot should be printed in the html file but also saved to disk as both png/pdf.
- Combine these scatter plots into one ggplot object using {patchwork} and print the combined plot in your .Rmd and .html. Save the combined plot to disk as a .pdf and .png in the "communications" folder.