Results Reproduction: Study 1

Jennifer Lin

11/26/2019

Contents

Introduction

The core of the thesis project is to understand the process in which the developers of the Moral Foundations Theory came to their conclusions about the differences between liberals and conservatives on moral grounds, and to apply this knowledge to other mechanisms of measuring morality to see if similar results can be achieved.

The notes in the replication of the results from Study 1 in Jesse Graham et al's 2009 paper that posits, from four different angles, that liberals and conservatives conceptualize morality in different domains.

Each of these commands have already been produced using an R script file, which can be located in the Git Repository under the `S1` folder housed under `ProjectImplicit`. The purpose of this document is to add notes for the author's own reference and for the general audience who wishes to reproduce this work.

Data for the Graham paper are posted on the Harvard Dataverse, and is freely available to the public. Here, I will use Study 1 data, which were both collected in the Project Implicit website.

For each dataset, the data was downloaded and cleaned in a R script file that will not be discussed in this file. Using those scripts, I extract pertinent variables and save to the directory as a `.csv` which will be sued in these analyses.

For the reproduction of the results in Study 1, the data cleaning file is `S1Management.R` and the original R script (reproduced in this document) is `Fig1.R`.

Before I begin, I will load all the packages that were used in the analyses in this section.

```
library(car)  #Recoding Data
library(dplyr)
```

```r
library(psych)  #Descriptive data
library(ggplot2)  #Create figures
library(GGally)
library("ggpubr")
library("reshape2")
library(scales)
library(lme4)  #lmer function
library(tidyverse)  #dplyr
library(lsr)  #EtaSquared
```

The goal of Study 1 is to understand moral relevance issues related to each of the five foundations. Here, the authors created a series of 20 questions that speak to whether or not an issue is relevant, in the eyes of the participant, to their judgment of morality. Participants respond to the set of questions on a scale of 1 (not relevant at all) to 7 (extremely relevant). Each of these items are distributed roughly evenly across the five foundations.

An attention check, whether someone believed in astrology, was introduced during the course of the survey to ensure participants were paying attention to the measures. Those who indicated the higher end of the scale were excluded from the analysis.

Figure 1 – Moral Relevance

To describe the data in the section, the authors created a figure where the average scores on each foundation were summarized across each level of the self-identified political ideology spectrum.

Below, I will comment on the reproduction process for Figure 1

```r
# Load in the data
morals <- read.csv("~/Desktop/Working/Moral-Psychology/ProjectImplicit/PI-study1.csv",
    header = TRUE, na.strings = c("", " ", "NA"))
```

Once the data is loaded, I will recode the data such that they align with the scales that are used in the paper. For the files, the responses are coded from 1-6, but in the paper, the responses are averaged on a scale of 0-5. As such, each variable will be recoded so that it matches the 0-5 scale.

```r
# Harm
morals$suffer <- morals$suffered1 - 1
morals$violence <- morals$violence1 - 1
morals$harmed <- morals$harmed1 - 1
morals$Harm <- rowMeans(morals[, c("suffer", "violence", "harmed")],
    na.rm = TRUE)

# Fairness
```

```r
morals$differently <- morals$differently1 - 1
morals$rights <- morals$rights1 - 1
morals$unfairly <- morals$unfairly1 - 1
morals$Fairness <- rowMeans(morals[, c("differently", "rights",
    "unfairly")], na.rm = TRUE)

# Ingroup
morals$betray <- morals$betray1 - 1
morals$friend <- morals$friend1 - 1
morals$loyalty <- morals$loyalty1 - 1
morals$Ingroup <- rowMeans(morals[, c("betray", "friend", "loyalty")],
    na.rm = TRUE)

# Authority
morals$rank <- morals$rank1 - 1
morals$duties <- morals$duties1 - 1
morals$respect <- morals$respect1 - 1
morals$Authority <- rowMeans(morals[, c("rank", "duties", "respect")],
    na.rm = TRUE)

# Purity
morals$disgust <- morals$disgust1 - 1
morals$purity <- morals$purity1 - 1
morals$unnatural <- morals$unnatural1 - 1
morals$Purity <- rowMeans(morals[, c("disgust", "purity", "unnatural")],
    na.rm = TRUE)
```

Now that each of the questions are grouped and recoded, foundations are created as their own variable and reflect the average score for each participant across the variables.

Next, I will recode political ideology and rename the levels on the variable to match that of the graphs that are used in the paper.

```r
table(morals$politics)  #Look at the data distribution
```

```
##
##  -3  -2  -1   0   1   2   3
## 262 440 162 348 117 110  30
```

```r
#-1 = Liberal, 1 = Conservative - Higher scores mean more liberal/conservative
morals$ideology <- as.character(as.integer(morals$politics))
```

In the graph, each of the levels on the ideology variable are labeled in character strings. Here, I will recode each of the variables to match those in the paper.

```r
morals$ideology <- recode(morals$ideology, `-3` = "Strongly Liberal")
morals$ideology <- recode(morals$ideology, `-2` = "Moderately Liberal")
```

```r
morals$ideology <- recode(morals$ideology, `-1` = "Slightly Liberal")
morals$ideology <- recode(morals$ideology, `0` = "Neutral")
morals$ideology <- recode(morals$ideology, `1` = "Slightly Conservative")
morals$ideology <- recode(morals$ideology, `2` = "Moderately Conservative")
morals$ideology <- recode(morals$ideology, `3` = "Strongly Conservative")
```

Now, let's look at the recoded ideology variable to see if it is done correctly

```r
table(morals$ideology)
```

```
##
## Moderately Conservative       Moderately Liberal                  Neutral
##                     110                      440                      348
##   Slightly Conservative         Slightly Liberal    Strongly Conservative
##                     117                      162                       30
##        Strongly Liberal
##                     262
```

In the original labeling, the higher the number, the more extreme one is to their respective side. Additionally, the negative numbers represent the liberals and the positives represent the conservatives. Here, the recode matches the original labels.

Now, with such recoding method, there is bound to be some implicit NA's from people who do not answer the questions. Here, we will clear these implicit values.

```r
library(forcats)
morals$ideology <- fct_explicit_na(morals$ideology, na_level = "NA")
table(morals$ideology)
```

```
##
## Moderately Conservative       Moderately Liberal                  Neutral
##                     110                      440                      348
##   Slightly Conservative         Slightly Liberal    Strongly Conservative
##                     117                      162                       30
##        Strongly Liberal                       NA
##                     262                       79
```

In addition, to graph the results, the labels need to be switched to factors rather than maintain the original character strings that we are currently working with. Next, I will convert the variables to factors and order them as desired, removing the NA's

```r
morals$ideology <- as.factor(as.character(morals$ideology))
morals$ideology <- factor(morals$ideology, levels = c("Strongly Liberal",
    "Moderately Liberal", "Slightly Liberal", "Neutral", "Slightly Conservative",
    "Moderately Conservative", "Strongly Conservative"))
```

There are many ways that R can use to aggregate values by another variable. The way I chose to do it is a bit messy but it gets the job done. Here, I use the `Aggregate()` command

and execite it across each foundation, stroing each foundation as its own data frame.

```
Harm <- aggregate(Harm ~ ideology, morals, mean, na.rm = TRUE)
Fairness <- aggregate(Fairness ~ ideology, morals, mean, na.rm = TRUE)
Ingroup <- aggregate(Ingroup ~ ideology, morals, mean, na.rm = TRUE)
Authority <- aggregate(Authority ~ ideology, morals, mean, na.rm = TRUE)
Purity <- aggregate(Purity ~ ideology, morals, mean, na.rm = TRUE)
```

Since this generages 5 unique data frames, we would need to merge it into one succinct frame in order to plot the data.

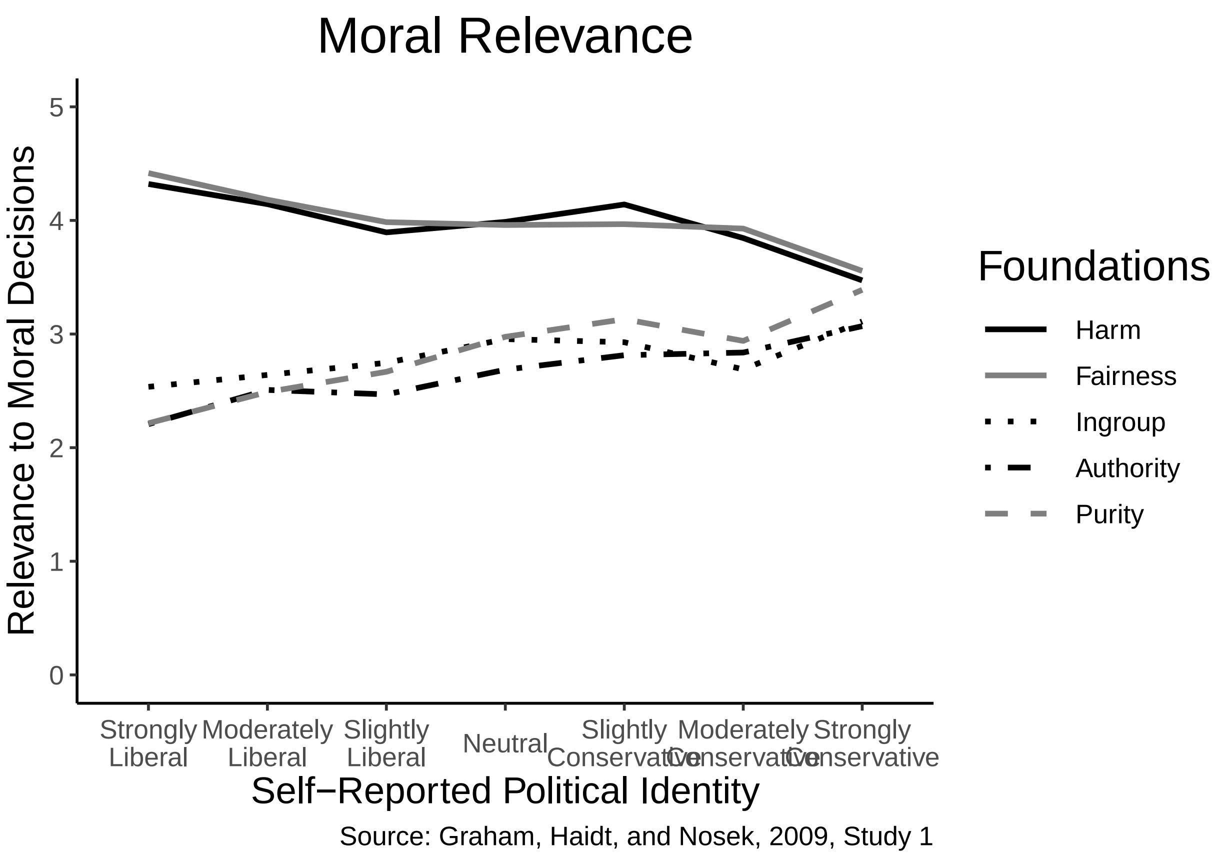
```
moral <- merge(Harm, Fairness, by.x = "ideology", by.y = "ideology",
    all.x = TRUE, all.y = TRUE)
moral <- merge(moral, Ingroup, by.x = "ideology", by.y = "ideology",
    all.x = TRUE, all.y = TRUE)
moral <- merge(moral, Authority, by.x = "ideology", by.y = "ideology",
    all.x = TRUE, all.y = TRUE)
moral <- merge(moral, Purity, by.x = "ideology", by.y = "ideology",
    all.x = TRUE, all.y = TRUE)
```

The output for this is also a bit messy such that the foundations are the column labels and the ideologies are the row labels. For the graph, we would need ideology to run as a variable in its own column, the foundation to also be grouped in it's own variable, and the aggregates as the third unique variable. As a result, we need to reshape the data using the `reshape` package.

```
mfq <- reshape2::melt(moral, id.var = "ideology")
```

Finally! All the work leads up to this point. We are ready to produce Figure 1 of the Graham et al study using the `ggplot2` package and other useful `gg-` packages loaded earlier.

```
ggplot(mfq, aes(x = ideology, y = value, group = variable)) +
    geom_line(aes(linetype = variable, color = variable), size = 1) +
    theme_classic() + scale_linetype_manual("Foundations", breaks = c("Harm",
    "Fairness", "Ingroup", "Authority", "Purity"), values = c(Harm = "solid",
    Fairness = "solid", Ingroup = "dotted", Authority = "dotdash",
    Purity = "dashed")) + scale_color_manual("Foundations", breaks = c("Harm",
    "Fairness", "Ingroup", "Authority", "Purity"), values = c(Harm = "black",
    Fairness = "grey50", Ingroup = "black", Authority = "black",
    Purity = "grey50")) + ggtitle("Moral Relevance") + xlab("Self-Reported Political Ide
    ylab("Relevance to Moral Decisions") + ylim(0, 5) + labs(caption = "Source: Graham,
    theme(text = element_text(size = 12, colour = "black"), axis.title = element_text(s
        colour = "black"), title = element_text(size = 16, colour = "black"),
        plot.caption = element_text(size = 10, color = "black"),
        axis.text.x = element_text(angle = 0, hjust = 0.5, vjust = 0.5),
        plot.title = element_text(hjust = 0.5), legend.key.width = unit(2,
            "line")) + scale_x_discrete(labels = wrap_format(10))
```



Repeated Measures General Linear Model

In the paper, the authors conducted a repeated measures GLM to see if the differences between aggregate individualizing versus aggregate binding foundations are different from one another and whether this effect was moderated by politics. To replicate the analysis, I build a linear model to demonstrate this effect.

First, I build an aggregate individualizing and binding foundation score.

```
# Individualizing and Binding Foundations Aggregate Variables
morals$indiv <- rowMeans(morals[, c("suffer", "violence", "harmed",
    "differently", "rights", "unfairly")], na.rm = TRUE)
morals$bind <- rowMeans(morals[, c("betray", "friend", "loyalty",
    "rank", "duties", "respect", "disgust", "purity", "unnatural")],
    na.rm = TRUE)
```

Next, I create a difference score between the individualizing and binding foundations. Note that if the scores are flipped, the ultimate result would not be affected.

```
morals$diffscore <- morals$indiv - morals$bind
```

I generate a linear model as follows

```r
diff.mode <- lm(diffscore ~ politics, data = morals)
summary(diff.mode)
```

```
##
## Call:
## lm(formula = diffscore ~ politics, data = morals)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.3130 -0.5450  0.0040  0.5563  3.1119
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.21170    0.02784   43.52   <2e-16 ***
## politics    -0.22548    0.01505  -14.98   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8332 on 1207 degrees of freedom
##   (339 observations deleted due to missingness)
## Multiple R-squared:  0.1568, Adjusted R-squared:  0.1561
## F-statistic: 224.5 on 1 and 1207 DF,  p-value: < 2.2e-16
```

```r
etaSquared(diff.mode)
```

```
##             eta.sq eta.sq.part
## politics 0.1568469   0.1568469
```