Moral Self-Concept and Charitable Giving: Pre-Analysis Plan

Justin Abraham and Nick Otis

November 7, 2016

Abstract

This document describes the pre-analysis plan for a randomized experiment examining the relationship between moral self-concept and charitable giving. Using the Amazon Mechanical Turk, respondents will complete an implicit association test (IAT), a questionnaire on support for redistributive polices, negative affect, self-esteem, omission bias; and have an opportunity to give a portion of their earnings to a charitable organization. A randomly assigned treatment group will be told that they posses strong implicit bias, while the control group will remain unaware of their results. The pre-analysis plan outlines our hypotheses, schedule of tasks, and empirical strategy. In order to guarantee transparency and bind ourselves from "fishing for results" we will pre-register the program producing our main results along with this pre-analysis plan.

1 Introduction

When people engage in charitable donation they are engaging in moral behavior. A number of recent theoretical developments have focused on understanding how people make moral decisions. Some authors emphasize the role of maintaining a positive moral identity in reference to some moral norm (Bénabou and Tirole 2011), while others stress the role of "social intuitions", gut feelings about moral actions that are later justified through post-hoc reasoning (Haidt 2001), or "pure altruism" (Jankowski 2002). However, little experimental research has explored in depth what factors causally influence moral decision making in the area of altruistic giving. Specifically, we seek to understand how an individual's perceived moral character (defined here as one's evaluation or conception of themselves as a moral person) is associated with their moral behavior, which serves the role of self-signaling that one is moral.

From this perspective, one of the drivers of moral decisions is the extent to which people feel as though they must maintain their perceived moral character in comparison to some moral norm or reference. Perceived moral character may depreciate over time if one does not repeatedly self-signal that they are a moral person through moral actions. When an individual incurs a loss in perceived moral character that brings them below what they believe the "moral norm" to be, they engage in moral actions until they have restored their perceived moral character to acceptable levels.

To investigate this mechanism, we conduct an online experiment on a widely used platform for experimental research, Amazon Mechanical Turk (MTurk). We test this maintenance model of moral character by presenting a randomly assigned group of respondents with information on their level of implicit racial bias, which acts as a moral self-signal, but should not influence charitable donation based on social intuitions or pure altruism. We then observe whether a hit to moral

stock caused by being informed of one's implicit bias influences charitable donation, a mechanism for restoring moral stock. Additionally, we will administer several scales to parse affective and conceptual mechanisms that influence decisions. In particular, we examine the effects on omission bias, by which harm incurred from inaction is systematically deflated (Ritov and Baron 1999).

To tie our hands from data mining for significant results, this document outlines our pre-analysis plan specifying the statistical models we will use to identify the causal effect of our experimental manipulation on charitable giving behavior. We will register this pre-analysis plan and scripts[1] for data analysis prior to data collection. We will pilot the survey instrument and script with subjects recruited from MTurk and exclude these observations from the main study.

2 Experimental Design

2.1 Sample

We will conduct our experiment on the MTurk platform. This platform is now commonly used by researchers to conduct online experiments as it provides a cheap and efficient way of recruiting respondents (Paolacci, Chandler, and Ipeirotis 2010). We will recruit workers on MTurk over the age of 18 and residing in the United States. We will screen for and exclude non-English speakers. Each respondent completing the survey and meeting the study's inclusion criteria receives a minimum compensation of $1. Respondents have the opportunity to earn up to $2 based on survey responses.

2.2 Schedule of tasks

The following outlines a schedule of tasks in the order completed by respondents after completing the consent form. At the conclusion of the study, the survey displays a debriefing statement and payment information.[2]

1. **Pre-treatment questionnaire**: This section collects a variety of pre-treatment respondent characteristics including gender, age, race, income, employment status, level of education, political attitude, and explicit racism.

2. **Implicit association test (IAT)**: The IAT is a sorting task that uses response time to measure associations with a target concept. This task consists of five blocks. The first block asks respondents to associate 16 randomly ordered words with "Good" and "Bad" attributes. The second block asks respondents to associate 12 randomly ordered faces with "Black" and "White". The third block asks respondents to associate words *and* faces with "Black/Good" and "White/Bad". The fourth block reverses the positions of the attributes in the first block and the fifth block uses "Black/Bad" and "White/Good" as attributes. Response latency will be calculated using the third and fifth blocks and following Greenwald, Nosek, and Banaji (2003). Half of our sample will be randomly assigned to receive their IAT results.

3. **Negative affect and self-esteem**: Immediately after the completion of the IAT, the respondent will answer four 6-point Likert questions measuring negative affect (afraid, upset, nervous, and distressed). We also include a Likert-type question on self-esteem.

[1] The program for data cleaning and analys is available here for download: `https://gist.githubusercontent.com/jrpabraham/51ac90c24785b52058455296dd5e911f/raw/f2d11cbc7b05f7210feab75f378df3a0d743b74c/MS_Pilot.do`

[2] Debriefing statements, and experimental tasks are included as part of the Appendix.

4. **Charitable giving**: Respondents are told they will receive $1 in addition to the fixed compensation advertised on MTurk. They are given the opportunity to donate any portion of this windfall to a charitable organization. The amount specified will be deducted from this windfall and donated to the organization at the conclusion of the study.

5. **Trolley problem**: Respondents are presented with the trolley problem thought experiment, first presented by Foot (1967). It asks readers to hypothetically choose between saving the lives of five people versus one person. Respondents first read the "passive" version of the thought experiment wherein the onlooker must pull a lever then read the "active" version which involves pushing a bystander.

6. **Support for redistribution**: Respondents are asked to indicate their support for redistributive attitudes using a 7-point Likert scale.

7. **Support for affirmative action**: Respondents are asked to indicate their support for affirmative action using a 7-point Likert scale.

2.3 Treatment

This study investigates the effect of moral self-concept on charitable giving. We experimentally manipulate moral self-concept by randomly assigning half of respondents into the treatment group to receive an interpretation of their IAT results, emphasizing participants implicit racial bias. The control group will receive confirmation of completion of the IAT without results. Revealing to respondents a potentially unacknowledged implicit bias acts as a negative shock to moral self-concept.

In addition to the IAT manipulation, we overlay a second stage of randomization when soliciting explicit racism. A random half of the sample will be asked "Are you racist?" while the other half will skip this item. We will explore heterogeneity of the treatment effect by being primed to explicitly consider racist attitudes and identity.

3 Empirical Strategy

3.1 Analysis of potential selection bias

To assess whether differential selection into treatment confounds our results, we estimate the following equation to examine the correlation between assignment and a vector of pre-treatment observables $\mathbf{X}_i$ described in Section 4.2.

$$T_i = \alpha_0 + \mathbf{X}_i'\omega + \varepsilon_i \tag{1}$$

3.2 Treatment effect

Our main specification is given by the following model.

$$Y_i = \beta_0 + \beta_1 T_i + \varepsilon_i \tag{2}$$

Y_i denotes the outcome of interest, further described in Section 4.1. T_i is an indicator for treatment assignment, and ε_i is the idiosyncratic error term. Thus, β_1 identifies the treatment effect of having been shown IAT results.

To improve precision, we also estimate Equation 3, which includes a full set of interactions with indicators of pre-treatment variables described in Section 4.2. Define a vector of demeaned indicators $\dot{\mathbf{X}}_i = \mathbf{X}_i - \bar{\mathbf{X}}_i$. We obtain the covariate-adjusted treatment effect estimate by estimating Equation 2 including $\dot{\mathbf{X}}_i$ as an additive term and as an interaction with the treatment indicator.

$$Y_i = \beta_0 + \beta_1 T_i + \dot{\mathbf{X}}_i'\gamma_0 + T_i\dot{\mathbf{X}}_i'\gamma_1 + \varepsilon_i \tag{3}$$

The set of indicators partitions our sample so that our estimate for β_1 remains unbiased for the average treatment effect (Lin 2013). We will report treatment effect estimates with and without covariate adjustment.

3.3 Heterogeneous treatment effect

We will test whether the treatment effect varies with pre-treatment characteristics. Pre-specified dimensions of heterogeneity are described in Section 4.3. To investigate heterogeneity by baseline characteristic W_i, we estimate a model that interacts receipt of transfers with W_i.

$$Y_i = \beta_0 + \beta_1 T_i + \beta_2 W_i + \gamma(T_i \times W_i) + \varepsilon_i \tag{4}$$

3.4 Exact test of the treatment effect

In addition to the large-sample approach outlined in Section 3.2, we leverage the randomized design of this study to perform exact tests of the cross-village treatment effect (Fisher 1935). Randomization inference allows us to test the Fisherian sharp null hypothesis that $y_i^{(t)} = y_i^{(c)}$ for every unit i.[3] To remain analogous with our main specification, we use the conventional Wald statistic from Equation 2 as our test statistic. We calculate exact p-values for this statistic under the null hypothesis using a Fisher permutation test. Specifically, we take 10,000 permutations of the treatment indicator and calculate the Wald statistic for each m^{th} permutation. The exact p-value is

$$\frac{1}{10,000}\sum_{m=1}^{10,000} \mathbb{1}\Big[\hat{\beta}_m'\hat{V}(\hat{\beta}_m)^{-1}\hat{\beta}_m \geq \hat{\beta}_{obs.}'\hat{V}(\hat{\beta}_{obs.})^{-1}\hat{\beta}_{obs.}\Big] \tag{5}$$

Unlike the model-based approach, randomization inference treats outcomes as fixed and the distribution of the estimate arising from the random assignment of the treatment. This method allows us to conduct valid inference on a finite sample without relying on asymptotic properties. Moreover, the exact test is uniformly more powerful and asymptotically identical to the standard t-test with a simple comparison of two groups (Lehmann and Romano 2006). We will report results from both approaches after data analysis.

[3] This is not equivalent to a null hypothesis of no average treatment effect.

4 Variables of interest

4.1 Outcome variables

To test our primary hypotheses, we analyze the treatment effect on the following outcomes.

1. Percent of windfall donated to charity
2. Support for affirmative action score
3. Support for redistribution score
4. Trolley problem choice

We also analyze the following outcomes as manipulation checks of the treatment.

1. Negative affect score
2. Self-esteem response

4.2 Control variables

The following pre-treatment variables will be used to test balance between treatment and control groups and will be included in some specifications for covariate adjustment.

1. English is respondent's primary language
2. Respondent is female
3. Respondent is over 30 years old
4. Respondent has some college/university education
5. Respondent lives in an urban locality
6. Respondent lives in suburban locality
7. Respondent earns at least $50,000 annually
8. Respondent is unemployed
9. Respondent identifies as African American
10. Respondent leans politically liberal
11. Respondent asked about explicit racism

4.3 Dimensions of heterogeneity

1. Respondent has some college/university education
2. Respondent earns at least $50,000 annually
3. Respondent is unemployed
4. Respondent identifies as African American
5. Respondent leans politically liberal
6. Respondent asked about explicit racism

References

Bénabou, Roland, and Jean Tirole. 2011. "Identity, morals, and taboos: Beliefs as assets." *The Quarterly Journal of Economics* 126 (2): 805–855.

Fisher, Ronald Aylmer. 1935. *The Design of Experiments*. Edinburgh: Oliver & Boyd. Includes index. Bibliography: p. 245.

Foot, Philippa. 1967. "The problem of abortion and the doctrine of double effect."

Greenwald, Anthony G., Brian A. Nosek, and Mahzarin R. Banaji. 2003. "Understanding and using the Implicit Association Test: I. An improved scoring algorithm." *Journal of Personality and Social Psychology* 85 (2): 197–216.

Haidt, Jonathan. 2001. "The emotional dog and its rational tail: a social intuitionist approach to moral judgment." *Psychological review* 108 (4): 814.

Jankowski, Richard. 2002. "Buying a Lottery Ticket to Help the Poor Altruism, Civic Duty, and Self-interest in the Decision to Vote." *Rationality and Society* 14 (1): 55–77.

Lehmann, Erich L., and Joseph P. Romano. 2006. *Testing Statistical Hypotheses*. Springer Science & Business Media.

Lin, Winston. 2013. "Agnostic notes on regression adjustments to experimental data: Reexamining Freedman's critique." *The Annals of Applied Statistics* 7 (1): 295–318 (March).

Paolacci, Gabriele, Jesse Chandler, and Panagiotis G. Ipeirotis. 2010. "Running experiments on amazon mechanical turk." *Judgment and Decision making* 5 (5): 411–419.

Ritov, Ilana, and Jonathan Baron. 1999. "Protected values and omission bias." *Organizational behavior and human decision processes* 79 (2): 79–94.

A Experimental tasks

A.1 Baseline questionnaire

Lan: What is your primary language (i.e., the one you speak most of the time)?
1 English
2 Spanish
3 Chinese

4 French
5 Arabic
6 Other [STRING]

Gen: What is your gender?
0 Male
1 Female

YeaBor: What year were you born?
[INTEGER]

Edu: Please indicate the highest level of education you have completed.
1 Primary School
2 High School or equivalent
3 Vocational/Technical School (2 year)
4 Some College College Graduate (4 year)
5 Master's Degree (MS)
6 Doctoral Degree (PhD)
7 Professional Degree (MD, JD, etc.)
8 Other

Loc: Which of the following best describes the area you live in?
1 Urban
2 Suburban
3 Rural

Pol: In general are you:
0 more of a conservative
1 more of a liberal

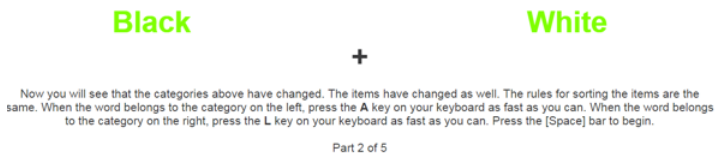
Inc: Please indicate your current household income in U.S. dollars
1 Under $10,000
2 $10,000 - $19,999
3 $20,000 - $29,999
4 $30,000 - $39,999
5 $40,000 - $49,999
6 $50,000 - $74,999
7 $75,000 - $99,999
8 $100,000 - $150,000
9 Over $150,000

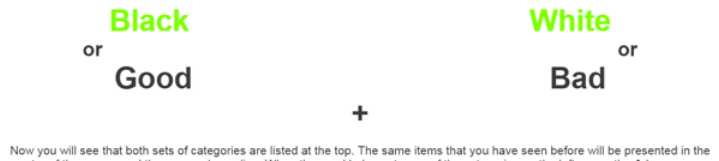
Emp: What best describes your employment status?
1 Full Time
2 Part Time
3 Retired

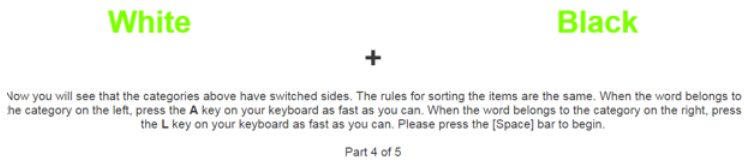
4 Unemployed

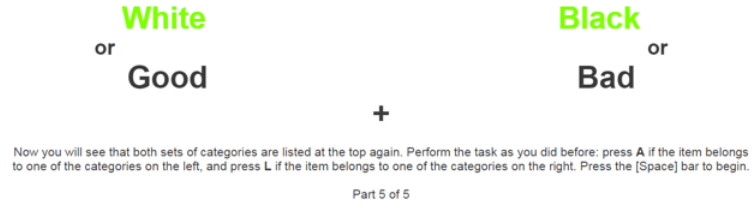
Rac: With which racial group do you most identify with?
1 Hispanic or Latino
2 Black or African American
3 American Indian or Alaska Native
4 White
5 Asian
6 Native Hawaiian or Pacific Islander

A.2 Implicit association test









Based on random assignment, respondents will receive one of the following two responses after completion of the IAT:

No information:

You have completed the Race Implicit Association Task. You may now proceed to the next question.

Implied racist:

You have completed the Race Implicit Association Task. Your data suggest that **you have implicit attitudes *against* African Americans** compared to White Americans. You may now proceed to the next question.

TreIat:
0 No information
1 Implied racist

A.3 Explicit racism

Based on random assignment, respondents will randomly be asked a question about explicit racism:

TreExpRac: Respondent is asked explicit racism question
0 No
1 Yes

ExpRac: Are you a racist?
0 No
1 Yes

A.4 Negative affect and self-esteem

This scale consists of a number of words that describe different feelings and emotions. Read each item and then list the number from the scale below next to each word. Indicate to what extent you feel this way right now, that is, at the present moment.

NegAfi: To what extent do you feel afraid?

NegUps: To what extent do you feel upset?

NegNer: To what extent do you feel nervous?

NegDis: To what extent do you feel distressed?

SelEst: On the whole, I am satisfied with myself?

A.5 Charitable giving

ChaDon: In addition to your participation payment, you will receive $1.00 bonus that you can choose to donate to GiveWell, an organization which evaluates charities on the basis of impact. Your donation will go toward GiveWell's top-rated charity. GiveWell collects no fees from donations. Whatever amount you give will be deducted from your account and donated directly to the charity at the end of the day and you will keep the remainder. For example, if you enter 50, have of this amount will go to you and the other half will be donated. What percentage of your earnings do you want to give? Please enter a number from 0-100.

A.6 Trolley problems

TroSwi: A runaway trolley is coming down the track. It is headed towards five people who cannot get out of its way. You can save the five people by throwing a switch and diverting the trolley down a siding. If you do, the trolley will collide with and kill a lone man standing on the siding. Do you throw the switch and divert the trolley?

0 Do nothing
1 Yes, divert the trolley

TroFat: A runaway trolley is coming down the track. It is headed towards five people who cannot get out of its way. If you push a nearby large man onto the tracks his bulk will stop the trolley before it hits the five, though the large man himself will be killed. Do you push the man and divert the trolley?

0 Do nothing
1 Yes, divert the trolley

A.7 Support for redistribution

For each of the following statements, indicate whether you strongly disagree, disagree, are neutral, agree, or strongly agree with the statement.

RedRes: As a country's wealth increases, more of its resources should be channeled to the poor.

RedBen: Giving to others usually benefits the givers as well.

RedObl: Those with more resources have more obligations toward their fellow human beings.

RedWel: It is beneficial for all to spend money on the public sector in education, housing, and healthcare.

RedExp: Those who are well off cannot be expected to take care of everyone else.

RedDep: Charitable organizations just create dependency among the recipients.

A.8 Support for affirmative action

For each of the following statements, indicate whether you strongly disagree, disagree, are neutral, agree, or strongly agree with the statement.

AffFai: Affirmative action for African Americans is unfair to White Americans.

AffEdu: Affirmative action in educations gives an opportunity to qualified African Americans who might not have had the chance without it.

AffHir: Affirmative action for African Americans may force employers to hire unqualified people.

AffCom: Affirmative action in the workplace for African Americans helps make sure that the American workforce and economy remain competitive.

B Debriefing statement

This study aims to investigate the relationship between implicit biases and pro-social behavior.
We first collected basic demographic information and participants support for different social policies. In order to measure implicit attitudes, participants conducted an implicit association test (IAT) which measures response times to different stimuli. The purpose of this task is to detect the strength of automatic associations between different objects or concepts of interest. This provides a measure of implicit attitudes regarding the object of interest. The study recorded participants true IAT results and asked about perceived performance in the task. Participants displayed results does not reflect participants actual results. Finally, we measured pro-social behavior by observing the proportion of study earnings each participant gave to a charitable organization.

Again, thank you for participating in this study. Please contact the investigators at the following e-mail addresses if you have any questions or concerns regarding this study.

Justin Abraham
justinra@princeton.edu

Nicholas Otis
notis@princeton.edu