

## **PRE-ANALYSIS PLAN**

### **Altruism and Policy: Experimental Evidence from Elite Civil Servants in Pakistan**

#### **1. Background**

This is a pre-analysis plan specifying the analysis we plan to perform for the evaluation of a training workshop on empathy in Pakistan with deputy ministers on their actual policy outcomes in year 1 and year 2 post-workshop as they perform their official duties. Specifically, we will use government audit documents to investigate spending decisions of ministers across 12 government departments (which represent the universe of their spending choices).

Our main hypothesis is that altruism training will increase spending on social policies (i.e. will raise ministers' budgetary recommendations to education and health) and reduce spending on self-oriented policies (i.e. budgetary recommendations for personal security and office maintenance).

This study experimentally evaluates three methods of cultivating prosocial behaviour (i) Utilitarian Value of Empathy (ii) Malleability of Empathy (iii) Joint Utilitarian and Malleability Treatment (iv) Placebo training.

This is an individual level randomised controlled trial with 213 deputy ministers with about 50 ministers per treatment arm and we evaluate impact on ministers' 12 policy outcomes. If we deviate from the methods set out in this document, we will provide clear justifications and wherever possible will also publish pre-plan analysis for comparison.

## **2. Introduction**

Prosociality—behavior that benefits others or society as a whole—is critical in contract enforcement, management of commons, public goods provision, establishing effective rule of law, efficient governance in societies and labour market success. This raises an important policy question: How can prosociality be cultivated? Beyond laboratory studies showing short-term malleability of prosocial behavior, there have been few field experiments to train prosociality effectively, especially in adults. A pioneering experiment found improvements in prosociality after an early childhood intervention, while recent experiments build on this and find improvements in prosociality from one year of mentoring of elementary school children and from a yearlong, three-hour-per-week curriculum designed to build social cohesion in schools. We explore an effective, scalable way to train prosociality among adults through the utilitarian value of empathy. We horse race among different schools of thought on cultivating prosociality and test Trivers’s reciprocal altruism against cultivating prosociality by emphasising the malleability of the self.

## **3. Research Questions**

Can training ministers in empathy impact their policy? Will it induce them to spend more on education and health of citizens? Will empathy training cause crowd-out of other policies?

## **4. Treatment Details**

In this study, we conducted a randomized evaluation with junior deputy ministers at a deputy minister’s training academy in Pakistan. The primary objective of the study was to investigate the effects of empathy training on policy making among junior ministers. The training facility integrated different methods of empathy training into their regular curriculum, motivated by the results of an earlier survey which found that a significant proportion of junior deputy ministers joined the civil service for personal gains rather than public service.

To speak to this issue, we randomized junior ministers into four training workshops, including one placebo training in macroeconomics.

The first training workshop focused on the utility of empathy in policy making, emphasizing the benefits of empathic behavior in organizational and individual performance. This was achieved through a combination of narratives and quantitative evidence from the private sector. In the second workshop, junior ministers were randomized into the malleability of empathy training, which presented the idea that empathy is a skill that can be developed and improved over time. This message was delivered through narratives of individuals who demonstrated growth in empathy and quantitative evidence from the private sector.

The third workshop combined the key messages from both the utility and malleability of empathy training, emphasizing both the benefits and the potential for improvement. In addition, we evaluated the impact of these three empathy training programs against the placebo training in macroeconomics.

This study consisted of four stages: Stage I involved a 30-minute recorded lecture followed by a writing exercise on the main lessons learned. Stage II consisted of a 2-hour live Zoom session, including a 10-minute structured discussion, a 50-minute lecture, and a 1-hour interactive activity session. Stage III involved a book assignment, cross-randomizing junior ministers to either receive an empathy or a placebo book, followed by a written assignment. Finally, in Stage IV, we investigate the impact of the treatments on policy decisions occurring one and two years post-treatment.

## 5. Empirical Specification

The impact of the training will be evaluated by comparing outcomes across groups in a simple regression framework. For each bureaucrat, the estimation equation is as follows:

$$Y_i = \alpha + \beta U_i + \gamma M_i + \delta UM_i + X_i \mu + \epsilon_i \quad (1)$$

where  $Y_i$  is respective policy outcome for deputy minister  $i$ ,  $U_i$  is a dummy equal to one if the deputy minister is assigned to the stand-alone utilitarian empathy treatment arm;  $M_i$  is a dummy variable equal to one if the deputy minister is assigned to the stand-alone malleability empathy treatment arm;  $UM_i$  is a dummy variable equal to one if the deputy minister is assigned to the joint

utilitarian and malleability treatment arm;  $X_i$  is a vector of individual-level controls. We cluster standard errors at the individual level since that is our level of randomization. In equation (1),  $\beta$  measures the effect of stand-alone utilitarian treatment;  $\gamma$  the effect of stand-alone malleability treatment; and  $\delta$  the effect of the joint treatment. The experiment with deputy ministers will also solicit demand for metrics training and control for this in all specifications. We will provide results with and without all available controls. We will estimate the parameters both with and without controlling for baseline characteristics such as written test scores, interview test scores, gender, birth in political capitals, asset ownership, income before joining civil service, age, education, foreign visits, and occupational group dummies. etc. We will test the null hypothesis that  $\beta = \gamma = \delta = 0$ , check for no differential attrition by treatment status, and cluster standard errors at the individual level in our analysis.

If we find evidence of differential attrition across treatment groups we will explicitly model the attrition (based on observable characteristics at baseline) and report estimates that are corrected for it, as well as bounds on the estimates where necessary.

We will assess the baseline balance of the endline sample by assessing the difference in means between treatment arms across baseline characteristics, accounting for clustering of standard errors and adjusting for the fact that we are testing for multiple hypotheses. We will use anderson-q values and List et al., (2019)'s familywise error rate correction (FWER) which extends the False Discovery Rate (FDR) method by using a bootstrapping approach, incorporating point-dependence structure of different treatments and controlling for the familywise error rate i.e., the probability of one or more false rejections. Specifically, we will apply a test that pools p-values across both outcomes and treatments in a single family i.e. all 12 policy outcomes will be pooled into a single family.

## 6. Outcomes Variables

The outcomes on policy decisions are for fiscal support or budgetary requests of deputy ministers from administrative data on their performance evaluation records.

*Data and Outcomes.* —Our data on policy is derived from performance audits of deputy ministers which contains information on their spending on various government departments. There are 12 government departments on which they can choose to recommend a budgetary allocation. They are:

- 1) health,
- 2) education,
- 3) personal security,
- 4) office budget,
- 5) IT,
- 6) development,
- 7) horticulture,
- 8) sports,
- 9) entertainment,
- 10) law and order,
- 11) climate encroachment,
- 12) disaster.

*Main Hypotheses.* —In our pre-analysis, we outline two main categories of policies that are most likely to be impacted by altruism training. For instance, we expect social policies, i.e. spending on health and education departments, to be positively impacted and selfish (self-oriented) policies, i.e. personal security and office budget, to be negatively affected (as they potentially crowd-out government funds) or unaffected (if government budget is not binding and there is no crowd-out). In the case of crowd-out, these selfish policies would be negatively impacted by increased value of altruism. Our main hypothesis, therefore, is that spending on social policies will increase and spending on selfish policies will decrease, while the other 8 policies will be unaffected.

*Robustness Analysis.*— The following robustness tests will be conducted at a minimum: check for balance over individual characteristics, test for differential attrition, and correction for multiple hypotheses testing.

*Hypotheses Tested.* — We will test the following main hypotheses:

H1: Utilitarian training impacts one of the 12 policies after one and two years of intervention. Utilitarian training increases the probability of empathy book choice.

H2: Utilitarian training has no impact on any of the 12 policy choices one and two years after intervention. Utilitarian training has no impact on probability of empathy book choice.

H3: Malleability training impacts one of the 12 policies after one and two years of intervention. Malleability training increases the probability of empathy book choice.

H4: Malleability training has no impact on any of the 12 policy choices after one and two years of intervention. Malleability training has no impact on probability of empathy book choice.

H5: Utilitarian + Malleability training impacts one of the 12 policies after one and two years of intervention. Utilitarian + Malleability training increases probability of empathy book choice.

H6: Utilitarian + Malleability training has no impact on any of the 12 policy choices after one and two years of intervention. Utilitarian + Malleability training has no impact on probability of empathy book choice.

**Causal mediation design:** 
$$Y_i = \alpha + \beta U_i \times \text{Empathy Book Assigned} + \gamma M_i \times \text{Empathy Book Assigned} + \delta U M_i \times \text{Empathy Book Assigned} + \lambda \text{Empathy Book Assigned} + \theta \text{Book Chosen} + X_i \mu + \epsilon_i \quad (2)$$

H7: Utilitarian, Malleability, and/or Utilitarian + Malleability training impacts one of the 12 policies after one and two years of intervention but only when assigned the empathy book.

H8: Utilitarian, Malleability, and/or Utilitarian + Malleability training impacts one of the 12 policies after one and two years of intervention but only when assigned the empathy book