Pre-Analysis Plan for

The effect of misinformation training on policymakers: Evidence from Pakistan

Description

In recent years, false and misleading news has proliferated online, especially on social media. The effects of this trend may be especially pernicious when policymakers themselves fall prey to untrustworthy information and amplify it. In this study, we test whether exposure to a training workshop on the dangers of misinformation and how to recognize it helps deputy ministers and tax officers in Pakistan to better distinguish between false or misleading news and legitimate news and reduces the extent to which they would share false or misleading news online.

Hypotheses

We will test the following hypotheses:
H1: Participants in the misinformation workshop will perceive misinformation as a more severe problem (H1a) and believe it is important to avoid spreading it (H1b).

H2: Participants in the misinformation workshop will better distinguish between false or misleading news and legitimate news in their behavioral sharing intentions (H2a) and accuracy ratings (H2b).

H3: Participants in the misinformation workshop will be less likely to select false or misleading news for inclusion in a government briefing.

H4: Participants in the misinformation workshop will be more likely to support government intervention to restrict false information online.

Statistical models to be used in the analysis

We will compute treatment effects via OLS with robust standard errors. We will use a lasso variable selection procedure to determine the set of prognostic covariates to include in each model (i.e., separately for each model below). If one or more levels are selected from a factor variable, we will include only the selected level(s) in the model. Below is the set of candidate variables that we will select from (as measured pretreatment) -- if not otherwise stated, these variables will be treated as continuous:
-highest degree indicators
-region or territory of origin indicators
-years in government service
-political interest
-conspiratorial predispositions
-trust and confidence participants in mass media
-how frequently participants use Twitter
-how frequently participants read political news on Twitter
-how frequently participants tweet or retweet political news on Twitter

-whether the participant or their staff uses or controls a Twitter account (indicator)
-political knowledge

Our models will also include pretreatment measures of the outcome variable when available and fixed effects for the type of official (tax officer, less senior deputy ministers, more senior deputy ministers).

All models will be estimated on midline data collected starting the day after the intervention and on endline data collected approximately two weeks later.

To test our hypotheses, we will estimate the following models corresponding to the variables defined above:

H1a: misinformation_severity = B0 + B1*treatment + B2*misinformation_severity_pretreatment + official type fixed effects + lasso-selected controls

H1b: avoid_spreading_misinformation = B0 + B1*treatment + B2*avoid_spreading_misinformation_pretreatment + official type fixed effects + lasso-selected controls

H2a: sharing_discernment = B0 + B1*treatment + B2*sharing_discernment_pretreatment + official type fixed effects + lasso-selected controls

H2b: accuracy_discernment = B0 + B1*treatment + official type fixed effects + lasso-selected controls

H3: briefing_misinformation = B0 + B1*treatment + official type fixed effects + lasso-selected controls

H4: government_intervention = B0 + B1*treatment + official type fixed effects + lasso-selected controls

Inference criteria

We will use the standard p<.05 criteria for rejecting null hypotheses but will also report whether results are significant at the p<.01 and p<.005 levels (two-sided).

Data exclusion

Data will be discarded for those who answer “No” for the consent question.

Missing data

“Don’t know” or blank responses will be considered missing data for outcomes. If our pre-specified covariate-adjusted models lead to more than 20% of observations being dropped, we will use multiple imputation. We will otherwise use listwise deletion.

Exploratory analysis and Heterogenous treatment effects

-We will estimate exploratory models testing for heterogeneous treatment effects by official type (tax officers, less senior deputy ministers, more senior deputy ministers) by interacting the treatment variable in the models above with indicators for the latter two official types.
-We will report treatment effects on sharing intention and accuracy ratings separately for legitimate and false or misleading news in addition to the discernment measures described above.
-We will compute and report summary statistics for our sample. We may also collect and may report response timing data as a proxy for respondent attention.
-The order of hypotheses and analyses in the final manuscript may be altered for expositional clarity.
-We may use the manipulation check variables reported above to estimate average treatment effects on the treated using two stage least squares if compliance rates are not sufficiently high.

Further details on the Study design

Experiment
A researcher randomly assigns treatments to study subjects, this includes field or lab experiments. This is also known as an intervention experiment and includes randomized controlled trials

Randomization
Participants in each cohort of Pakistani civil servants (tax officers, less senior deputy ministers, and more senior deputy ministers) will be independently randomized with probability=.5 at the individual level by the researchers into a treatment or placebo group. Those in the treatment

group will be assigned as part of their academy training to attend the treatment presentation and those in the placebo group will be assigned as part of their academy training to attend the placebo presentation.

Data collection procedures
Participants will be deputy ministers and tax officers in the Pakistani civil service who are taking part in training as part of their career development. The survey will be terminated for those who decline to consent. We expect to recruit approximately 800 participants.

The training groups and workshop details are below:
-Treatment - tax officers: July 5, 2022
-Placebo - tax officers: July 6, 2022
-Treatment - deputy ministers (less senior): July 6, 2022
-Placebo - deputy ministers (less senior): July 7, 2022
-Treatment - deputy ministers (more senior): July 27, 2022
-Placebo - deputy ministers (more senior): July 28, 2022

Surveys will be distributed to participants approximately 1-4 days before the workshop they attend, the day after the workshop they attend, and 15 days after the workshop they attend.

Sample size

We expect to recruit approximately 800 participants but the exact total may vary.

Sample size rationale

This is the maximum sample size we can recruit.

Stopping rule

We will collect data for all participants who take part in the training sessions listed above.

Variables

Manipulated variables

-An indicator that takes the value of 1 if the respondent is randomly assigned to the treatment condition and 0 otherwise (treatment).

Outcome variables

Prior to the treatment, we will measure the following:

… highest degree (indicator variables omitting bachelor's)

… region or territory of origin (indicator variables omitting Punjab)

… years in government service (1=6-10 years…5=26+ years)

… political interest on a five-point scale where 1 = “Not at all interested” and 5 = “Extremely interested”

… conspiratorial predispositions on a five-point scale where 1 = “Strongly disagree” that “Much of our lives are being controlled by plots hatched in secret places” and 5=“Strongly agree”

… how much trust and confidence participants have in mass media (newspaper, TV, radio, and social media) on a 4 point scale where 1 = “Not at all” and 4 = “A great deal”.

… how frequently participants use Twitter on a 7-point scale where 1 = “Never” and 7 = “Daily”

… how frequently participants read political news on Twitter where 1 = “Never” and 7 = “Daily”

… how frequently participants tweet or retweet political news on Twitter where 1 = “Never” and 7 = “Daily”

… whether the participant or their staff uses or controls a Twitter account (0=none, 1=1 or more)

… political knowledge on a three-point scale measuring the number of correct answers to three knowledge questions

After the treatment, we will measure the following:

… treatment compliance where 1=in the treatment group and answered both questions abotu the presentation correctly and 0 otherwise

… recommending including a false or misleading news headline in a government briefing (where 1=selected a false or misleading headline for inclusion and 0 otherwise)

Indices

We will calculate the following mean values at the respondent level separately before and after the treatment:

… mean sharing intention for legitimate news (where 1 = “Extremely unlikely” and 6=“Extremely likely”)

… mean sharing intention for false or misleading news (where 1 = “Extremely unlikely” and 6=“Extremely likely”)

… mean difference in sharing intention between legitimate and false or misleading news (the difference of the means above calculated at the respondent level)

… a composite measure of support for government action to restrict misinformation (the mean response to “Governments should take steps to restrict false information online, even if it limits people from freely publishing or accessing information” where 1 = “Strongly disagree” and 5 = “Strongly agree” and “People’s freedom to publish and access information should be protected, even if it means false information can also be published” where 1 = “Strongly agree” and 5 = “Strongly disagree”; we will split these and analyze separately if they do not correlate at r=.6 or greater)

… a composite measure of the perceived importance of not sharing misinformation (the mean response to the perceived importance of only sharing accurate information where 1 = “Not at all important” and 5 = “Extremely important” and agreement that people in Pakistan’s Civil Service need to avoid spreading misinformation where 1 = “Strongly disagree” and 5 = “Strongly disagree”; we will split these and analyze separately if they do not correlate at r=.6 or greater)

… a composite measure of the perceived severity of misinformation (the mean response to the perceived seriousness of misinformation in Pakistan where 1 = “Not at all serious” and 5 = “Extremely serious”, self-reported concern about misinformation in Pakistan where 1 = “Not at all concerned” and 5 = “Extremely concerned”, and agreement that people in the civil service often spread misinformation where 1 = “Strongly disagree” and 5 = “Strongly agree”; we will split these and analyze separately if they do not clearly load on a single factor in an exploratory factor analysis)

We will calculate the following mean values at the respondent level after the treatment:

… mean perceived accuracy of legitimate news (where 1 = “Not at all accurate” and 4 = “Very accurate”

… mean sharing intention for false or misleading news (where 1 = “Not at all accurate” and 4 = “Very accurate”

… mean difference in perceived accuracy between legitimate and false or misleading news (the difference of the means above calculated at the respondent level)