How does a fraud scandal impact trust in science? (Pre-Analysis Plan)*

Leon Eliezer† Luigi Garzon‡ Nícolas Goulart§

April 23, 2024

Abstract

This study addresses how unethical behavior in scientific research impacts trust in science, contributing to a literature on the consequences of corruption and unethical behavior. We propose an informational delivery experiment with high school students in Brazil, where we present them with a summary of a suspected fraud scandal in behavioral science. We divide treated subjects into a "fraud" arm, where they only learn about the fraud accusations, and an "accountability" arm, where they learn about the investigation and punishment processes involved; this allows us to tell apart the effects of learning about cheating from catching cheaters. We measure the effect on beliefs in science, the use of scientific evidence to update world views, real life attitudes towards science, and spillovers to other fields of research. The results of this experiment can enlighten academic institutions about the consequences of fraud, and guide public communication with respect to informing about accountability.

Keywords: fraud, corruption, belief formation, trust in science, information delivery experiment

JEL codes: D73, D83, D84, I23

Study pre-registration:

Proposed timeline: The post-treatment data for this study will be obtained by the end of April 2024.

*We thank Nicolás Ajzenman and Bruno Ferman for their suggestions and for their excellent advising throughout this entire project. We thank Yuehao Bai, Adrian Blattner, Daniel Da Mata, Vitor Possebom, João Pugliese, Vlasta Rasocha, Pedro Carregã Sant'Anna, Otávio Tecchio, and the seminar audience at FGV-EESP for their valuable comments. We thank Bladimir Carrillo, Daniel Da Mata and Fernanda Estevan for their comments on the scientific evidence presented to students in the survey. We thank everyone at OBECON for doing their best to allow this project to happen. We thank FGV-EESP for financing the operation.

†São Paulo School of Economics - FGV. leonveliezer@gmail.com

‡São Paulo School of Economics - FGV. luigigarzon0@gmail.com

§São Paulo School of Economics - FGV. nicolas.gou@hotmail.com

1 Introduction

Scientific research aims to produce evidence and create knowledge, not only to enhance our understanding of the world, but also with the goal of informing public opinion and government policy, hopefully making social choices better informed. However, for this goal to be attained, it is essential that voters and policymakers believe the scientific process and its results to be truthful. A body of recent evidence shows that trust in science has been in decline in the United States and elsewhere, specially since the COVID-19 pandemic, with a sharp politicization of science in society (Hamilton and Safford (2021), Gauchat (2012)). Among potential reasons for this decline are that epidemics may reduce trust in science (Eichengreen et al., 2021), that scientists and scientific institutions have been taking political stances (Motta (2018), Lee (2020), Sullivan (2020)), or a perception of reduced quality of published research, ranging from a reproducibility crisis (Ioannidis, 2005) to the flagging of unethical behavior, such as plagiarism (Hartocollis, 2024) and results manipulation (Lewis-Kraus, 2023).

This study uses an information provision experiment to understand to what extent unethical behavior in academia harms scientific credibility, and how communication about accountability can mitigate this effect. We exploit a suspected fraud scandal involving two major researchers in behavioral sciences, Francesca Gino, from Harvard Business School and Dan Ariely, from Duke. After four of Gino's papers were targeted with evidence of potential fraud (Simohnson et al. (2023a), Simohnson et al. (2023b), Simohnson et al. (2023c), Simohnson et al. (2023d)), she was put on administrative leave. Ariely also had a study suspected of fraud by the same blog two years prior (Simohnson et al., 2021a), which resurfaced after the company who provided the data for said study supported the claims of data tempering (Fountain et al., 2023).[1] Beyond understanding the consequences of these particular events, this setting allows us to understand the effects of unethical behavior more broadly, including plagiarism, cheating or corruption.

In our setting, the fraud was uncovered by a blog of independent investigators, who had the support of the scientific community in taking their case to the authorities and pushing for accountability (Simohnson et al., 2021b). This allows us to answer an important question: whether holding cheaters accountable for their unethical behavior can mitigate any negative effects on credibility.

[1] See Lewis-Kraus (2023) for further detail on these events.

By randomly assigning one treatment arm with the information that said researchers allegedly committed fraud (henceforth, "fraud" arm) and one treatment arm with information about the alleged fraud and about the investigations that led to said discovery ("accountability" arm), we are able to understand whether accountability has a positive effect on credibility. Crucially, we investigate both the importance of accountability against unethical behavior within institutions, and the importance of communication about accountability, focusing on the process that allows us to know about misdeeds instead of the misdeeds themselves.

Our study takes place at the Brazilian Economics Olympiad (*Olimpíada Brasileira de Economia*, henceforth OBECON), a knowledge competition for students at high school or first year of college. This demographic is particularly interesting because it consists of dedicated students who are interested in economics and other social sciences. Beyond the relevance of their beliefs as citizens and voters, these students could change their future educational and labor market decisions, choosing not to pursue scientific research in the long run; furthermore, perceptions formed at the impressionable years could have a long-lasting effect (Krosnick and Alwin, 1989).

The intervention consists of presenting treated students with a clear summary of the fraud scandal. To distinguish the effects of learning about fraudulent practice from those of learning about punishment of fraudsters, our setting has two treatment arms: a "fraud" arm, where students are simply informed about the occurrence of the fraud scandal, and an "accountability" arm, where students are also informed about the investigations which gathered evidence of manipulation, the institutional process of punishing the cheaters, and the support of the scientific community for these investigations. Participants answer a pre-treatment survey where we collect their beliefs and characteristics, and read the intervention text before answering a post-treatment survey.

We collect outcomes at different stages of the experiment. We directly ask participants' beliefs about credibility of science, scientific fraud and accountability in science, before and after treatment. We also collect participants' opinions over sensitive topics, show them scientific evidence regarding these topics and collect opinions again, to measure belief updating. After the treatment, we elicit participants' willingness to engage with science-related actions, such as reading and sharing scientific evidence, and interest in science-related books and classes. We also vary these outcomes across different areas of scientific knowledge to capture spillovers between fields, and we offer participants the information to read more about the fraud scandal, to understand demand for information.

Just after the intervention, treated students are asked if they would like to read more about the fraud scandal. We then test for the presence of motivated beliefs and confirmation bias: our design allows us to capture whether participants who were previously more trustful of science or hold certain political views are more or less likely to read more about the scandal. We believe this answers an important question in the context of a fraud scandal and other credibility crises: who is most inclined to engage with this information? If individuals refuse to engage with new information when it conflicts with their current world view, this poses a challenge for communicators, even beyond science-related topics.

On the pre-treatment survey, participants give their opinions on sensitive, policy-relevant, real-world topics. On the post-treatment survey, they are shown scientific evidence regarding these topics and give their opinions again, so that we can measure how they use scientific evidence to update their beliefs. Participants can also click on links to these scientific articles, copy the link to share the articles, or click a button to share a post on 𝕏/Twitter about the articles; we track these actions in the survey.

From previous results in the existing literature, we might expect asymmetrical responses regarding belief updating and motivated beliefs. In Zhang (2023), after learning that Nature supported Biden, Trump supporters decrease their trust in Nature by a far larger amount than Biden supporters increase theirs, and demand less scientific information after treatment. Bursztyn et al. (2023) find that liberals are more likely than conservatives to support policy opposed to their political views when faced with scientific evidence. In our setting, we might expect the views of "*bolsonaristas*" to be less sensitive to scientific evidence. Regarding motivated beliefs, on the other hand, as far as conservatives are more distrustful of science, it could be that these groups have a greater tendency to "read more" about the treatment, demanding more negative information about science.

At the end of the post-treatment survey, students participate on a lottery to win a book, and elicit their three favorite books among a set offered to them. These books include books by the scientists accused of fraud, other books on diverse science-related topics, and non-science books. Some time after the intervention, participants are given the opportunity to subscribe to a behavioral economics class, and answer a third follow-up survey. This way, we are able to measure participants' interest in science and engagement in science-related activities, as well as non-immediate beliefs.

Broadly, our research is related to the literature on belief formation and belief updating (see

Benjamin (2019) for a review on nonstandard belief formation). More specifically, we relate to a literature that studies whether people use scientific evidence to update their beliefs. Different branches of literature present evidence that people are not fully Bayesian and their process of belief updating is biased, such as the literature on motivated beliefs (Eil and Rao (2011), Zimmermann (2020), Bénabou and Tirole (2003)) and willful ignorance (Grossman and van der Weele, 2016); there is evidence that even policymakers are subject to such biases (Banuri et al., 2019). Gentzkow et al. (2023) study how individuals may be biased towards searching for information they agree with, and Bursztyn et al. (2022) provide evidence that they are more likely to do so even when information is important and the misinformation is harmful for themselves.

In contrast, there is also evidence that individuals and policymakers take factual information into account when updating their beliefs and forming their opinions, and even more so regarding scientific evidence (Grigorieff et al. (2020), Haaland and Roth (2020), Hjort et al. (2021), Bursztyn et al. (2023)). One consequence of this behavior which has been explored by researchers and policymakers is how individuals react to misinformation. There is a branch of literature in social psychology addressing how individuals receive unfounded information (namely, "bullshit", henceforth BS), showing that BS is more well received when it is "pseudo-profound" or tries to pass as scientific (Pennycook et al. (2015), Evans et al. (2020), West and Bergstrom (2021), see Iacobucci and De Cicco (2022) for a review). Rafkin et al. (2021) show that when official communication provides wrong guidance, the public becomes more distrustful of future recommendations. Many experiments in economics have shown that providing individuals with accurate information and fact-checking can induce more correct beliefs and be an important tool for policymakers, including by changing the beliefs of politicians and bureaucrats themselves (Bowles et al. (2020), Enríquez et al. (2023), Bowles et al. (2023), Bursztyn et al. (2022), Rogger and Somani (2023)). Hjort et al. (2021) show that informing politicians with evidence about the effect of a policy makes them more likely to adopt it.

Our research is also inserted in a branch of literature that studies the credibility of science and the scientific community. Gauchat (2012) shows that trust in science has decreased in the US, specially among conservatives. There is also evidence that, when researchers and academic institutions take political stances, they harm the credibility of the scientific community, spilling over to other scientists and organizations (Motta (2018), Kotcher et al. (2017), Zhang (2023)).

Eichengreen et al. (2021) show that individuals who lived through epidemics are more distrustful or science and less likely to comply with health-related policies. In line with the existing research, we are interesting in understanding what can harm the reputation of the scientific community; our novel contribution lies in investigating the reputational effects of fraud scandals. Beyond that, we also investigate whether these effects also discourage young students from pursuing careers in science, which could harm scientific production itself in the long run.

In addition, our research is connected with the literature on the behavioral effects of unveiling corruption scandals. Ajzenman (2021) shows that revealing corruption by local politicians causes children to cheat more on tests, by changing their perception of how honesty and cheating are prevalent in the world. Both Barr and Serra (2010) and Gächter and Schulz (2016) show that young students from more corrupt countries are more likely to engage in dishonesty and corruption. These studies are similar to our paper in showing behavioral responses to corruption by young students in particular. Beyond that, other research also investigates the behavioral effects of unveiling corruption, such as Fisman and Miguel (2007) and d'Adda et al. (2017). Similarly to our paper, the latter experimentally varies the participant's perception of corruption. We contribute to this literature in two important ways. First, by investigating the effects of misconduct in academia, instead of public administration. Second, our design allows us to understand how accountability for unethical behavior can mitigate the consequences of said behavior, which is not explored by many papers in the literature about corruption scandals.

Finally, we relate to a strand of literature which studies the effect of anti-corruption messages. In particular, one strand of literature argues that corruption contains a component of "self-fulfilling prophecy": Corbacho et al. (2016) show that anti-corruption messages can increase participants' willingness to bribe officers, and Cheeseman and Peifer (2022) show that anti-corruption messages may backfire for individuals with a high prior perception of corruption. We contribute to this literature by learning about the effects of different contents of anti-corruption messages, specifically understanding the difference between communicating about corruption and communicating about accountability.

This plan proceeds as following. Section 2 details the experimental design. Section 3 describes how we collect data. Section 4 describes our empirical approach to test hypotheses and analyse data. Section 5 details our outcomes of interest and the hypotheses regarding them. Section

6 details how we will analyze heterogeneous effects to further investigate our research question. Section 7 details robustness checks we include in our design. Section 8 indicates how we interpret the different results we might obtain. Section 9 concludes the plan.

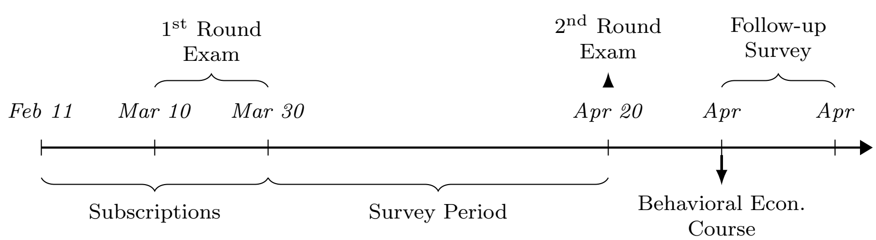
2 Research Design

2.1 Experimental Setting

The intervention will take place with the students of the Brazilian Economics Olympiad (*Olimpíada Brasileira de Economia*, or OBECON). OBECON is an economics competition among high-schoolers which happens every year since 2018. Thousands of students participate in the first round of exams, which takes place in an online test. The best ranked students then may classify to the second and third rounds, win medals and even represent Brazil at the International Economics Olympiad, providing further academic and professional opportunities later on.

Participants who consent with participating in the experiment are contacted and the experiment proceeds according to Figure 1 and described below:

Figure 1: Experiment Timeline



(i) After subscriptions to OBECON are closed and the first round of exams is done, consenting participants will be sent a baseline survey. Reminders will be sent for subjects who did not answer a first time, in order to increase participation.

(ii) A couple of weeks after that, participants are sent the post-treatment survey, which contains the intervention text and follow-up questions. We send the post-treatment survey before grades of the first stage are released, in order to minimize attrition from participants who didn't pass to the next stage.

(iii) Finally, we then send invitations for subjects to participate in a behavioural economics course a couple of weeks after the post-treatment survey, and send a third follow-up survey to collect belief changes which are not immediately after the survey.

2.2 Intervention

The intervention consists of reporting the suspected fraud scandals to the treated groups in the beginning of the post-treatment survey. One treatment arm, the "fraud" arm, will only receive information detailing the fraud accusations; another treatment arm, the "accountability" arm, will receive this message with a punishment framing, describing the institutional reactions to the scandal and the existing mechanisms in academia to prevent fraud and punish those who commit it. The control arm will be an active control: they will be informed about the story of Behavioral Economics, as well as the names of current researchers, including Gino an Ariely. We include an active control in order to separate the effect of learning about the fraud scandal *per se* from any potential effect of making salient behavioral economics or the researchers in question.

2.3 Randomization Design

One major concern in an information delivery experiment is that control participants learn about the information by interacting with treated participants, generating contamination bias. We consider that the main risk of contamination arises when students who go to the same school are assigned to different treatment arms, therefore talking about the experiment (potentially before one of them took the survey).

To assess the extent of this risk, we analyzed the distribution of participants between different schools in Figure 2, which comes from the potential participants in the subscription process. While a great share of participants come from schools with only one participant or few participants, there are relevant exceptions, where some schools have over 100 or 200 participants. Due to this concentration, to avoid risking contamination among these participants, we perform the randomization at the school level.

Our sample will consist of those individuals that properly completed the baseline survey. With this sample, we employ a matched-tuples design (Bai et al. (2023b)), *i.e.*, given a set $\mathcal{D} = \{1, \ldots, |\mathcal{D}|\}$ of possible treatments status, we match the schools into blocks of $|\mathcal{D}|$, and, within each block, we randomly assign all individuals in one school to each $d \in \mathcal{D}$. More specifically, we perform a matched-triplets design, since in our case $\mathcal{D} = \{1, 2, 3\}$, where $d = 1$ is the control arm and $d = 2, 3$ are the treatment arms. In order to match schools into groups of three, we perform the following blocking algorithm:

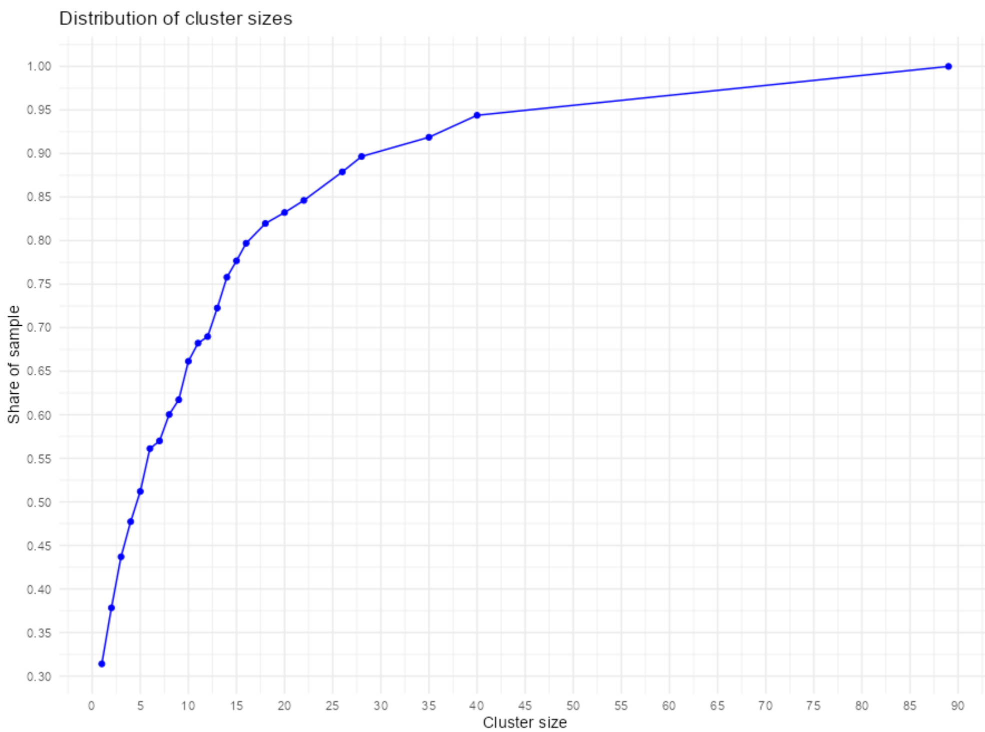
1. Group individual data at the school level and calculate the school mean for each variable.
2. Calculate the Mahalanobis distance between the clusters using both baseline outcomes and baseline covariates.
3. Pair the units such that it minimizes the sum of the Mahalanobis distances.
4. Select the 2/3 of the pairs that has the lower Mahalanobis distances.
5. Calculate the mean of both baseline outcomes and baseline covariates for each of those 2/3 pairs.
6. Using the mean of the outcomes and covariates, repeat steps 1 and 2 with the remaining clusters to get blocks of three units.
7. Perform a balance test. If there are striking differences between cluster size within the same triplet, which is not expected, we anticipate the possibility to modify the randomization algorithm, as to ensure we match clusters of similar size.

This is an adaptation of the "pairs of pairs" algorithm in Bai (2022). With this method, we ensure that the units within each block are close to each other in terms of baseline covariates, and that the treatment groups are balanced in terms of baseline covariates.

The baseline covariates used for matching will be cluster size, pre-treatment levels of trust in science, pre-treatment levels of interest in academic activities, pre-treatment opinions on sensitive topics, gender and political orientation. The randomization will occur after the end of the first round of OBECON and collection of baseline data.

Importantly, we only ensure balance within the subsample of participants who answered the baseline survey. Our final sample, however, are participants who answered both the baseline and endline survey, so dropouts may generate imbalances, even if by chance. In a balance test, we will be able to verify any potential imbalances.

Figure 2: Distribution of potential participants by cluster size



2.4 Power Calculation

Power calculations were made considering a significance level of 5% and a power of 80%. We compare our results to the benchmark of 0.15 standard deviations suggested by Haaland et al. (2023).

A number of 4830 participants in 1171 finished the first round OBECON exam and thus were contacted to participate in the experiment. Among these participants, around 1600 opted in, distributed among 664 schools. We expect this number to decrease until the endline survey. The distribution of school size among current opt-ins is shown in Figure 2. We expect a similar proportion of clusters to students to hold within the final sample.

We are interested in power for two main specifications. The first one, which is more powerful, aggregates both treatment arms, so treated proportion is $p = \frac{2}{3}$ and uses the entire sample N and a number of clusters N_g. The second one, which is less powerful but allows us to separate treatment arms, compares each treatment arm to control separately, so treated proportion is $p = \frac{1}{2}$ but each comparison uses $\frac{2}{3}N_g$ clusters, and approximately $\frac{2}{3}N$ individuals.

Following Eldridge et al. (2006), we correct our power calculations for heterogeneity in cluster size. We use the following formula to calculate the Minimum Detectable Effect (MDE):

$$\frac{MDE}{\sigma} = (t_{\frac{\alpha}{2}} + t_{\kappa})\sqrt{\frac{1}{Np(1-p)}}\sqrt{DE}$$

Where DE is the design effect:

$$DE = 1 + \rho(\overline{m}(CV^2 + 1) - 1)$$

ρ is the intra-cluster correlation coefficient, and CV is the cluster variance coefficient, defined as the standard deviation of cluster sizes to the mean cluster size. That is:

$$CV := \frac{1}{N_g - 1}\sqrt{\sum_{i=1}^{N_g}(m_i - \overline{m})^2}$$

m_i is the size of cluster i and $\overline{m} := \frac{1}{N_g}\sum_{i=1}^{N_g} m_i$.

We also take into account multiple hypotheses testing, as we will explicit further in Section 4.2. In order to make power calculations, we correct the significance level as $\alpha = 1 - (1 - 0.05)^{\frac{1}{h}}$, where h is the number of hypotheses inside a family. In our most conservative correction we use $h = 5$.

Below are our obtained results for power, varying with endline retention, intra-cluster correlation ρ and correction for multiple hypothesis testing. These results consider the most powerful specification, aggregating both treated arms.

Table 1: Power Calculations

		$\rho = 0$	$\rho = 0.05$	$\rho = 0.1$	$\rho = 0.25$	$\rho = 0.5$	$\rho = 1$
80% Retention ($N = 1269$)	$h = 1$	0.17	0.24	0.29	0.41	0.56	0.77
	$h = 5$	0.20	0.29	0.35	0.50	0.68	0.93
70% Retention ($N = 1110$)	$h = 1$	0.18	0.25	0.31	0.43	0.60	0.82
	$h = 5$	0.22	0.31	0.38	0.53	0.72	1.00

We focus on the specifications with smaller levels of intra-cluster correlation, since our choice for block randomization comes from the risk of contamination, not because we believe there are relevant common shocks at the cluster level. With this in mind, our study is well-powered, slightly above what would be considered ideal by the benchmark of Haaland et al. (2023).

3 Data Collection

We work with two main sources of data: data shared by the participants in our surveys and data shared by OBECON. Data shared by the participants is obtained from three different surveys: the baseline survey, the endline survey and the follow-up survey.

3.1 Baseline Survey

On this survey, after obtaining consent, we elicit participants' political opinions, their prior beliefs about science and level of trust in science, their self-perceived attitudes towards academic activities, and their opinions about socially sensitive issues. We know their answers, the time they spent on each page and whether they passed an attention check.

3.2 Endline Survey

On this survey, we show participants the intervention text and give them the opportunity to read more about the treatment if they belong to a treated arm. We elicit again beliefs in science, trust in science and self-perceived attitudes towards science. Then, we show participants a piece of scientific evidence regarding each socially sensitive issue they were asked on baseline.

3.3 Follow-up Survey

On this survey we elicit participants' beliefs in science, trust in science and self-perceived attitudes towards science a third time to capture the behavior of these variables some time after the treatment.

3.4 OBECON Data

We have access to data on information about students by our partner institution, OBECON. This includes the students' school of origin, social and demographic information asked on the subscription forms and participants' grades on the OBECON exams and their progress on the competition. Importantly, OBECON keep track of whether participants were caught cheating in the second round of exams, and raise red flags for potential cheaters in the first round of exams (if they spent too much time on the page, or participants from the same school who submitted the exam at the same time, for example).

4 Empirical Analysis

4.1 Estimation

For our main parameter of interest, the treatment effect for any treatment arm, we estimate equation (1) via OLS:

$$Y_{is} = \alpha + \sum_{d=2}^{3} \beta_d \mathbb{1}\{D_{is} = d\} + \mathbf{x'}_{is}\boldsymbol{\delta} + +\kappa_s \varepsilon_{is} \tag{1}$$

Where Y_{is} is the outcome of interest, for individual i in triplet s, D_{is} is the treatment status of participant i in triplet s, $\mathbf{x}_{is}$ is a vector of baseline covariates used in stratification, κ_s is a triplet fixed effect, and ε_{is} is an error term. We will also estimate specifications which aggregate both treated arms, fraud and accountability, to increase statistical power.

We interpret our parameters as Intention-To-Treat effects (ITTs). This is because many students could already have the information of the fraud scandal prior to the intervention, and because some students can simply answer the questions without reading.

4.2 Inference

We will follow the recommendation of Bai et al. (2023b) for inference in a matched tuples design, and perform a t-test with cluster-robust standard errors at the triplet level. These can be potentially too conservative with a large number of treatment arms, which we do not believe is our case. Furthermore, because of the uncertainty regarding the final sample size and the possibility that the final sample becomes smaller than desired, we reserve the possibility of performing inference via permutation tests to have an inference method that is valid in finite samples. In this case, we will take cluster-robust standard errors into account when performing the permutations.

Our main approach to deal with multiple hypothesis testing is to aggregate outcomes within the same family into an index, as to turn multiple answers into one single outcome. Nevertheless, we are interested in many families of outcomes. For this reason, it is important to pre-register a correction for multiple hypothesis testing. We plan to report p-values adjusted by the False Discovery Rate using the two-step procedure described in Benjamini et al. (2006). However, we appreciate that the Multiple Hypothesis Testing literature is evolving and that other new methods may show themselves to be clear improvements, in which case we will then report a deviation from this plan. In a similar fashion, we appreciate that p-value adjustments for MHT might potentially increase the rate of Type II errors and underpower the design, so we will report results both with and without these adjustments, taking this into account when interpreting our results.

4.3 Attrition

In our setting, participants are only exposed to treatment if they begin answering the Endline Survey. Thus, we can assume that dropping out before opening the survey is independent from treatment assignment. The risk of differential attrition, therefore, comes from incomplete answers at the Endline Survey.

We define an attrition dummy, A_i, to be equal to 1 if the individual dropped the Endline Survey mid-response. Ex-post, we do not expect to have attrition correlated with neither treatment nor control groups, so it will be equivalent to a sample size reduction. We will test for the presence of differential attrition with the following regression, conditional on those who clicked to view the Endline Survey:

$$A_i = \alpha + \sum_{d=2}^{3} \beta_d \mathbb{1}\{D_i = d\} + \epsilon_i \tag{2}$$

One common recommendation is to drop attritted triplets in a matched triplets scenario. Bai et al. (2023a) shows that dropping triplets do not recover the average treatment effect, but it potentially recovers a convex weighted average of conditional average treatment effects. However, Ferman and Ponczek (2017) show that this approach may lead to biased estimates, because observed attrition is a poor estimator of attrition probability when strata size is not large. Therefore, we will follow their recommendation and will not drop triplets with attritters, but simply include triplet fixed effects in the regression.

As for outcomes collected in a longer term, the risk of differential attrition is higher. We will perform the same test as in regression (2), where A_i is a dummy that states whether the person answered the follow-up survey. If attrition is too high, we may perform partial identification as in Lee (2009).

Furthermore, if attrition is high, we may track a random subsample of attritters in treated and control groups in order to test if they are significantly different in observables from non-attritters.

5 Outcomes

We are interested in understanding the effect of informing about a fraud scandal in the following families of outcomes: participants' self-reported trust towards science, their use of scientific evidence to update beliefs, and their real-life actions towards science. The table below describes the division between families of outcomes.

Table 2: Families of Outcomes

	Primary outcomes
Family 0. Informativeness of treatment	Knew about scandal; passed treatment quiz
Family I. Motivated demand for information	1. Choice to "read more" by political ideology 2. Choice to "read more" by baseline trust in science
Family II. Self-reported trust and attitudes toward science	1. Index of trust in science 2. Index of academic future
Family III. Use of evidence to update beliefs	3. Index of belief updating (aggregated and divided by category)
Family IV. Actions towards science	6. Choice of scientific books (aggregated and divided by category) 7. Clicked, copied or shared scientific evidence 8. Subscribed for behavioral econ. course

To test hypotheses regarding these outcomes in Equation (1), we first must determine what is the outcome variable, Y_i. In our preferred specification, we will attribute a scale of 1 to 4 to the responses, and standardize these values by the control mean and standard deviations. Because the cardinal values given to each response are arbitrary, this does not allow for a clear interpretation for the point estimates, but the sign of the estimator shows the direction in which people update their beliefs.

To deal with the problem of multiple hypothesis testing, we will construct a summary index of standardized outcome variables within the same family, weighing by the inverse of the covariance between outcomes, as in Anderson (2008). This approach allows us to test whether there is a general effect on that family of outcomes and eliminates the problem of multiple testing between outcomes in the same family.

Beyond that, we will also estimate two alternative specifications. One where we run a separate regression with a dummy for each point on the response scale as the outcome variable, for full transparency on the treatment effects on the distribution of answers, and one where we use as outcome variable the direction in which the participant changes their answer between baseline and endline.

5.1 Informativeness of Treatment

Before testing the effects of learning about fraud, we are interested in understanding whether participants absorb the information presented to them - that is, whether they learn about the fraud scandal after reading the intervention texts.

In the endline survey, just after the intervention text, participants are asked: (i) whether they knew about the information presented to them prior to the treatment, and (ii) whether a set of statements related to the text are true or false. This way, we can check how many participants report the fraud scandal as being new to them, and how many participants correctly report the fraud scandal as being true after the intervention.

We hypothesize the following regarding informativeness of treatment:

Hypothesis A.1 *Treated participants did not know about the fraud scandal prior to the experiment.*

Hypothesis A.2 *Treated participants will correctly acknowledge that Gino and Ariely were accused of fraud after the intervention.*

5.2 Motivated Demand for Information

After being shown the intervention text, participants on treated arms are given the opportunity to read more about the intervention. We use this to test a hypotheses of motivated demand for information in the context of the fraud scandal. The theory regarding motivated beliefs suggests that participants who were more distrustful of science on baseline are more interested in learning about a scandal of scientific fraud. In the context of this experiment, we are interested in learning whether there is difference in demand for information in two dimensions: participants' self-reported baseline trust in science, and participants' political position.

Hypothesis B.1 *Participants with more pro-science belies on baseline are less willing to acquire negative information about a fraud scandal in science.*

Hypothesis B.2 *Right-wing participants are more willing to acquire negative information about a fraud scandal in science.*

5.3 Beliefs in Science

Our first set of outcomes where we are interested in analyzing treatment effects are participants' self-reported levels of trust in science. Both in baseline and endline, we ask participants how much they agree with a statement about science, and how much they view themselves pursuing academic activities. Answers are collected on a 4-point response scale. Table 3 details these outcomes.

Table 3: Beliefs in Science Statements

Index	**Outcome**	**Sentence**	**Scale**
Index of academic future	Pursue major in Economics	Do you plan on pursuing an Economics major in college?	Certainly Not Probably Not Probably Yes Certainly Yes
	Work in academic activities	Do you see yourself working with academic activities in the future?	
	Interest in current research	Do you see yourself reading an article about recent scientific research?	
Index of trust in science	Overall trust in science	The scientific consensus about an issue is generally right	Totally Disagree Partially Disagree Partially Agree Totally Agree
	Beliefs about unethical behavior	Researchers often act unethically to manipulate their research results.	
	Beliefs about accountability	Scientific institutions can successfully prevent, identify and punish unethical behavior	

Our hypotheses regarding beliefs in science are:

Hypothesis C.1 *Being exposed to the information of fraud reduces subjects' trust in science.*

Hypothesis C.2 *Being exposed to the information of fraud reduces subjects' willingness to engage in academic activities.*

Hypothesis C.3 *Being informed about accountability in academic institutions mitigates this effect.*

To analyze how these effects last over some time after the treatment, we will send participants a third wave of surveys to collect their longer-term beliefs.

5.4 Using Evidence to Update Beliefs

Our second family of outcomes is how much individuals update beliefs when presented with scientific evidence. In baseline, individuals are asked how much they agree with each of the statements on Table 4. These topics are divided into those with evidence by Behavioral Economics, Economics or Natural Sciences. The fact that each evidence presented to them comes from each of those fields is made clear to the participants.

We do not intend to present evidence on these topics as if there is an unequivocal scientific consensus over said topic. Our approach is simply to show a piece of evidence whose conclusion sends a signal for participants to update beliefs in one direction.

Participants answer all affirmatives on a 4-point response scale similar to the "Disagree/Agree" scale presented in Table 3. We flip answers when necessary so that a positive value means agreeing with the evidence presented. We standardize each variable by control mean on endline and by sample mean on baseline.

For each statement s, we define their belief updating $U_{i,s} := Z^1_{i,s} - Z^0_{i,s}$ as the difference between participant i's endline and baseline standardized answers. The index of belief updating, our primary outcome, is just the sum of $U_{i,s}$ across s. As alternative specifications, we will also use as outcomes a dummy for whether the participant's answer updated in the direction implied by the evidence, and a fully saturated model where we interact the treatment with each level of prior beliefs.

We are also interested in decomposing all these outcomes separately by category and by topic, in which case we will adjust confidence intervals by Multiple Hypothesis Testing.

Our hypotheses regarding belief updating are:

Hypothesis D.1 *Subjects exposed to the information of fraud update less their beliefs when faced with scientific evidence, relative to the control group.*

Table 4: Opinions on Sensitive Topics

Category	Topic	Sentence	Paper
Behavioral Economics	**Payments based on performance**	Paying workers according to performance makes them more productive.	Ariely et al. (2009)
	Blood donation	Giving financial incentives for donating blood increases total donations.	Mellström and Johannesson (2008)
Economics	**Reelection**	The possibility of reelection is harmful for the political setting.	Ferraz and Finan (2011)
	Affirmative action	Quotas in public universities select less prepared students.	Oliveira et al. (2024)*
Natural Sciences	**Nuclear energy**	Nuclear energy is harmful for climate compared to other sources.	IEA (2022)
	GM organisms	Consuming transgenic food is bad for human health.	NASEM (2016)

*:For this topic, students are linked to a VoxDev article summarizing the paper: VoxDev (2024)

Hypothesis D.2 *Being informed about accountability in academic institutions mitigates this effect.*

5.5 Actions Towards Science

Beyond treatment effects over beliefs in science and belief updating, we are interested in whether learning about a fraud scandal makes participants change their actions regarding scientific-related activities. We analyze the following outcomes:

- **Choosing books on a lottery.** As a way of incentivizing the survey, those who answer the survey until the end participate on a lottery of free books. Participants rank their top 3 choice of books, among a pre-selected list presented to them in the survey. In this choice, we know whether participants are or not interested in science-related books. We will use as outcomes a dummy for science-related books, a dummy for each category of book, and a dummy for each book. In the specifications with a dummy for each book, we will report confidence intervals adjusted for multiple hypothesis testing.

- **Subscribing to a Behavioral Economics course.** A couple of weeks after the treatment, we offer participants the opportunity to sign up for a free Behavioral Economics course taught by the researchers. We will use as outcome whether the participant signed up or not, and whether the participant showed up or not (conditional on being selected).

- **Reading and sharing scientific evidence.** For each of the topics listed on Table 4, participants are given the opportunity to *click* on the link to the paper, to *copy* said link, and to *share the paper* on Twitter/𝕏. We collect said outcomes and aggregate them in an *index* of willingness to share scientific evidence.

- **Cheating.** Although cheating is unprecedented, we will know whether participants were involved in cheating in the subsequent rounds of the OBECON exams, in case they are approved. Therefore, we pre-register our interest in analyzing this outcome if it materializes with a relevant frequency.

Formally, our hypotheses regarding these outcomes are:

Hypothesis E.1 *Subjects exposed to the information of fraud are less likely to engage in academic-related activities.*

Hypothesis E.2 *Subjects exposed to the information of fraud are less likely to choose science-related books.*

Hypothesis E.3 *Subjects exposed to the information of fraud are less likely to read and share scientific evidence presented do them.*

Hypothesis E.4 *Subjects exposed to the information of fraud are more likely to cheat in an academic competition.*

5.6 Spillovers to Other Fields

Our next set of hypotheses is related to whether a fraud scandal by specific researchers in a specific field spills over to other fields.

Hypothesis F.1 *The effects of being informed about a fraud scandal in behavioral sciences spill over to other researchers and other areas of scientific knowledge.*

We test this hypothesis both in the book lottery and in the part of the survey where they are presented with scientific evidence of some topic. Both the categories of the books and the categories of the scientific evidence presented to them are divided in the following topics:

- Research by Gino/Ariely;
- Research in Behavioral Science;
- Research in Economics/Social Sciences;
- Research in Natural Sciences;
- Not scientific-related (in the case of the books).

6 Heterogeneity Analysis

To estimate heterogeneous effects along a given dimension, we include the cross-product terms in the regression. Let H be the number of possible values the heterogeneity dimension of interest, H_{is}, can attain. Then, we estimate equation 3 by OLS:

$$\begin{aligned} Y_{is} = \alpha + \sum_{d=2}^{3} \beta_d \mathbb{1}\{D_{is} = d\} + \sum_{h=2}^{H} \theta_h \mathbb{1}\{H_{is} = h\} + \sum_{d=2}^{3} \sum_{h=2}^{H} \delta_{dh} \mathbb{1}\{D_{is} = d, H_{is} = h\} \\ + \mathbf{x}'_{is} \boldsymbol{\delta} + \mu M_{is} + \kappa_s + \omega_{is} \end{aligned} \tag{3}$$

Where M_{is} is a dummy variable which indicates whether the participant chose to read more about the treatment (if the participant is in the control group, $M_{is} = 0$ always), and ω_{is} is an error term. We control for M_{is} because participants with different values of H_{is} might have different propensities to read more. Not controlling for M_{is} would confound the heterogeneity due to different treatment effects between groups from heterogeneity caused by increasing treatment intensity when participants choose to read more.

6.1 Treatment intensity

We are interested in understanding heterogeneous effects in different dimensions. First, we are interested in testing whether treatment effects are stronger when treatment is more intense:

Hypothesis G.1 *Treatment effects are larger for participants who spent more time at the treatment page.*

Hypothesis G.2 *Treatment effects are different for participants who chose to read more about the fraud scandal.*

Regarding G.1, we are not able to disentangle whether the heterogeneity comes from the extra information obtained or from a selection effect of participants who chose to read more. Nevertheless, this allows us to capture heterogeneity in a group that either (i) had more exposition to the information, or (ii) is particularly more interested in the information.

6.2 Heterogeneity by individual characteristics

We are interested in testing heterogeneous effects along a few individual characteristics of the participants, in order to understand what groups in our population react more to the fraud scandal.

Hypothesis H.1 *Treatment effects are different between women and men.*

Hypothesis H.2 *Treatment effects are heterogeneous students with different political views.*

Hypothesis H.3 *Treatment effects are heterogeneous among students with different test scores on the OBECON exam.*

Hypothesis H.4 *Treatment effects are heterogeneous among students who got red flags for potentially cheating on the exam.*

Our interest in analysing heterogeneous effects by gender in H.1 stems from the context of the fraud allegations. First, our intervention focuses on two main suspects, a man (Ariely) and a woman (Gino); as of this paper's publication date, Gino underwent investigation and was placed on administrative leave, while Ariely was not (Harvard (2024), Duke (2024)). Second, as part of her defense, Gino accused Harvard of gender discrimination. Potentially, male and female participants could perceive a gender discrimination component in fraud accusations, which can lead to a different interpretation of the information received.

Our interest in analysing heterogeneous effects among students with different political views in H.2 stems from existing results where right and left wing participants react differently to scientific evidence, such as Bursztyn et al. (2023), Zhang (2023). We go beyond left and right and also investigate heterogeneous effects among more specific political ideologies.

Understanding the heterogeneity in test scores as in H.3 is crucial because this paper aims to highlight how exposure to a fraud scandal could lead young students to drift away from pursuing an academic career, depriving the world from a positive externality which could possibly be created by their research. Because top ranked OBECON students are outliers with great potential to pursue research, heterogeneity on this dimension helps us assess the scale of this problem.

Finally, in H.4, in we are interested in heterogeneity by cheating red flags. Since the first round of the exam is take-home, examiners keep track of potential cheaters with orange and red flags: for example, students from the same school who turned in their tests at the same time, or students who kept the test page open for a long time. Students who view themselves as potential cheaters can react differently to the fact that cheaters are successful in the academic world; it is possible that being exposed to the fraud scandal encourages possible cheaters to pursue an academic career.

6.3 Heterogeneity by prior beliefs

As is usual in experiments that measure belief updating, we are interested in heterogeneity by prior beliefs. Furthermore, because we ask participants how certain they are of their answers, we are interested in heterogeneity by prior certainty.

Hypothesis I.1 *Treatment effects are different for participants with higher/lower baseline levels of trust on science.*

Hypothesis I.2 *Participants more certain of their priors are less likely to change their beliefs after treatment.*

7 Robustness

7.1 Variations of the intervention text

One identification concern is that, upon finding a statistically significant difference between treatment arms, the effect comes from the change in choice of words or other characteristics of the intervention texts, and not from the difference in information we would like to identify. For this reason, within each treatment arm, we present four slightly different versions of the same intervention text, but with minor alterations to wording and phrasing. The alterations are done by ChatGPT and verified by us so that the underlying message of the interventions do not change.

7.2 Attention checks

We include attention checks in the surveys to keep only valid answers. Both in baseline and endline, there are multiple choice questions which ask participants questions with trivial answers. Furthermore, the questions that measure whether the intervention text is informative can also work as attention checks. In our preferred specification, we exclude those who did not pass the attention checks.

7.3 Experimenter Demand Effects

One risk to which we are attentive are Experimenter Demand Effects that could be differential between treated and control participants. One approach we use to deal with this is to ask participants whether they think the survey intends to promote a particular political view, or a particular pro-science/anti-science view. Furthermore, we ask them to give the survey a rating overall. With this, we aim to provide evidence that any kind of experimenter demand generated by the survey is not differential between treatment arms.

8 Interpreting Results

We might initially expect that high school students, upon realizing that even academia is susceptible to fraudulent practices, could become less inclined to believe in scientific evidence. This issue is concerning because students need to absorb information based on scientific evidence as part of their daily learning. However, it is important to note that students generally understand that the scientific method is the most objective tool at our disposal for uncovering truths, despite the presence of a few instances of research misconduct. Therefore, if researchers have incentives to commit fraud, which could risk discrediting the entire scientific community, policymakers could focus on enhancing people's ability to question their initial thoughts and encouraging them to critically evaluate the content they encounter during information browsing (Iacobucci and De Cicco, 2022).

Besides, social media is becoming an increasingly widely used channel for disseminating journal articles which creates head-to-head competition against news stories (West and Bergstrom, 2021). For instance, the issue on the suspected fraud discussions was mainly discussed on $\mathbb{X}$ (formerly, Twitter) and blogs (Simohnson et al. (2023a), Simohnson et al. (2023b), Simohnson et al. (2023c), Simohnson et al. (2023d)). As a result, students might face difficulties in distinguishing well-founded scientific studies from sensationalized news stories, influenced by a similar framing effect. Furthermore, due to the propagation of fake news on the internet, the increasing competition between scientific information and sensationalism could further erode the credibility of science, thereby intensifying the mentioned effect.

Moreover, skeptical and agnostic students often engage with content that aligns with their existing beliefs or emotions and such events might amplify sensational aspects of scientific fraud stories to capture attention, potentially reinforcing confirmation bias among students (Douglas et al., 2019). Consequently, students might feel more inclined to remain inflexible in their existing beliefs, under the assumption that every article could potentially harbor errors or fraudulent information. We analyse this effect by testing for the presence of motivated beliefs.

For this reason, it is important to distinguish between the effects for the fraud and accountability arms. If we find that communicating participants about accountability in academic institutions makes participants more trustful of science or helps mitigate a negative effect, this implies that scientific rigor works to some extent in the goal of producing and transmitting credible knowledge. One direct implication of this interpretation is that, for science to maintain some credibility, academic institutions must work for catching and punishing cheaters.

One important feature of this experiment is that we measure effects on beliefs and outcomes some time after the intervention. This way, we are able to tell if the effect of the information is only immediate. Moreover, if we find no effects on the longer term outcome of subscribing to the behavioral economics class, we can test if treatment does not affect beliefs after some time, or if beliefs remain different but simply do not revert into actions for this specific outcome.

Another important feature of this experiment is that we measure effects for different areas of scientific knowledge. Therefore, we do not only document effects on trust and engagement for science in general, but also separately for hard sciences, social sciences, behavioral sciences and scientific work produced by those involved in the fraud scandal. As a consequence, even if we find no effect of the intervention on participants' beliefs and actions, our results would still have important implications. If participants still believe in scientific evidence and are interested in books produced by the researchers accused of fraud, this implies that the reputational cost of committing fraud is low among the general public. This means that if academic institutions want to create high costs for cheaters, the institutional and monetary punishments must provide large incentives on their own.

Finally, one differential aspect of our study is that, beyond the "demand-side" consequences that are normally addressed in research about trust in science, we also address the "supply-side" consequences. By measuring whether participants are less likely to study related fields or engage in scientific research, we also document whether distrust in science draws potential researchers away from research, damaging scientific production in the long run. If these effects are larger for students with high academic potential, the positive externality deprived from the world by drawing away these students is even higher.

9 Conclusion

This plan details an information delivery experiment to assess the effects of publicizing a suspected fraud scandal in behavioral economics. We examine effects on young students' beliefs in science, how they use scientific evidence to update their beliefs, and whether they are willing to engage in academic and research related activities. Not only are we interested in understanding whether suspected fraud affects belief formation and how this translates into actions, but we also want to understand if this effect is restricted to behavioral sciences or whether it spills over to different fields of research.

One innovative feature of this study is that, beyond investigating how suspected fraud affects science through credibility, we investigate the possibility that it harms science in the long run through a "scientific production" effect, by driving away young motivated students who could pursue a career in research. For this reason, we believe that an interesting extension to this experiment would be to evaluate the same effect in different populations. One could be interested in pursuing if students from other programs, such as the Physics Olympiad or Chemistry Olympiad would also be affected by this treatment. We could also question if college students, graduate students, or junior professors, who should have a deeper knowledge of science and academia and have already gone through their impressionable years, are also sensitive to this information treatment.

The evidence provided by this experiment contributes to the literature branch about how the reputation of the scientific community impacts individuals' beliefs and actions, and how that can be explored in shaping public communication and public policy. It is also relevant for the scientific community and academic institutions when measuring their actions against fraud and other kinds of unethical misconduct.

References

Ajzenman, N. (2021). The power of example: Corruption spurs corruption. *American Economic Journal: Applied Economics*, 13(2):230–57.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American Statistical Association*, 103(484):1481–1495.

Ariely, D., Gneezy, U., Loewenstein, G., and Mazar, N. (2009). Large Stakes and Big Mistakes. *The Review of Economic Studies*, 76(2):451–469.

Bai, Y. (2022). Optimality of matched-pair designs in randomized controlled trials. *American Economic Review*, 112(12):3911–40.

Bai, Y., Hsieh, M. H., Liu, J., and Tabord-Meehan, M. (2023a). Revisiting the analysis of matched-pair and stratified experiments in the presence of attrition.

Bai, Y., Li, J., and Tabord-Meehan, M. (2023b). Inference for matched tuples and fully blocked factorial designs. *Working Paper.*

Banuri, S., Dercon, S., and Gauri, V. (2019). Biased Policy Professionals. *The World Bank Economic Review*, 33(2):310–327.

Barr, A. and Serra, D. (2010). Corruption and culture: An experimental analysis. *Journal of Public Economics*, 94(11):862–869.

Benjamin, D. J. (2019). Chapter 2 - errors in probabilistic reasoning and judgment biases. In Bernheim, B. D., DellaVigna, S., and Laibson, D., editors, *Handbook of Behavioral Economics - Foundations and Applications 2*, volume 2 of *Handbook of Behavioral Economics: Applications and Foundations 1*, pages 69–186. North-Holland.

Benjamini, Y., Krieger, A. M., and Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, 93(3):491–507.

Bowles, J., Croke, K., Larreguy, H., Liu, S., and Marshall, J. (2023). Sustaining exposure to fact-checks: Misinformation discernment, media consumption, and its political implications. *Mimeo*.

Bowles, J., Larreguy, H., and Liu, S. (2020). Countering misinformation via whatsapp: Preliminary evidence from the covid-19 pandemic in zimbabwe. *PLOS ONE*, 15(10):1–11.

Bursztyn, L., Egorov, G., Haaland, I., Rao, A., and Roth, C. (2023). Justifying Dissent*. *The Quarterly Journal of Economics*, 138(3):1403–1451.

Bursztyn, L., Rao, A., Roth, C., and Yanagizawa-Drott, D. (2022). Opinions as Facts. *The Review of Economic Studies*, 90(4):1832–1864.

Bénabou, R. and Tirole, J. (2003). Intrinsic and Extrinsic Motivation. *The Review of Economic Studies*, 70(3):489–520.

Cheeseman, N. and Peifer, C. (2022). The curse of good intentions: Why anticorruption messaging can encourage bribery. *American Political Science Review*, 116(3):1081–1095.

Corbacho, A., Gingerich, D. W., Oliveros, V., and Ruiz-Vega, M. (2016). Corruption as a self-fulfilling prophecy: Evidence from a survey experiment in costa rica. *American Journal of Political Science*, 60(4):1077–1092.

Douglas, K. M., Uscinski, J. E., Sutton, R. M., Cichocka, A., Nefes, T., Ang, C. S., and Deravi, F. (2019). Understanding conspiracy theories. *Political Psychology*.

Duke (2024). Dan ariely (duke fuqua school of business). https://www.fuqua.duke.edu/faculty/dan-ariely. Accessed: Jan 07 2024.

d'Adda, G., Darai, D., Pavanini, N., and Weber, R. A. (2017). Do Leaders Affect Ethical Conduct? *Journal of the European Economic Association*, 15(6):1177–1213.

Eichengreen, B., Aksoy, C. G., and Saka, O. (2021). Revenge of the experts: Will covid-19 renew or diminish public trust in science? *Journal of Public Economics*, 193:104343.

Eil, D. and Rao, J. M. (2011). The good news-bad news effect: Asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics*, 3(2):114–38.

Eldridge, S. M., Ashby, D., and Kerry, S. (2006). Sample size for cluster randomized trials: effect of coefficient of variation of cluster size and analysis method. *International Journal of Epidemiology*, 35(5):1292–1300.

Enríquez, J. R., Larreguy, H., Marshall, J., and Simpser, A. (2023). Mass political information on social media: Facebook ads, electoral saturation and electoral accountability in mexico. *Mimeo*.

Evans, A., Sleegers, W., and Mlakar, (2020). Individual differences in receptivity to scientific bullshit. *Judgment and Decision Making*, 15(3):401–412.

Ferman, B. and Ponczek, V. (2017). Should we drop covariate cells with attrition problems? *Mimeo*.

Ferraz, C. and Finan, F. (2011). Electoral accountability and corruption: Evidence from the audits of local governments. *American Economic Review*, 101(4):1274–1311.

Fisman, R. and Miguel, E. (2007). Corruption, norms, and legal enforcement: Evidence from diplomatic parking tickets. *Journal of Political Economy*, 115(6):1020–1048.

Fountain, N., Guo, J., Romer, K., and Peaslee, E. (2023). Fabricated data in research about honesty. you can't make this stuff up. or, can you?

Gauchat, G. (2012). Politicization of science in the public sphere: A study of public trust in the united states, 1974 to 2010. *American Sociological Review*, 77(2):167–187.

Gentzkow, M., Wong, M. B., and Zhang, A. T. (2023). Ideological bias and trust in information sources. *Mimeo*.

Grigorieff, A., Roth, C., and Ubfal, D. (2020). Does information change attitudes toward immigrants? *Demography*, 57(3):1117–1143.

Grossman, Z. and van der Weele, J. J. (2016). Self-Image and Willful Ignorance in Social

Decisions. *Journal of the European Economic Association*, 15(1):173–217.

Gächter, S. and Schulz, J. (2016). Intrinsic honesty and the prevalence of rule violations across societies. *Nature*.

Haaland, I. and Roth, C. (2020). Labor market concerns and support for immigration. *Journal of Public Economics*, 191:104256.

Haaland, I., Roth, C., and Wohlfart, J. (2023). Designing information provision experiments. *Journal of Economic Literature*, 61(1):3–40.

Hamilton, L. C. and Safford, T. G. (2021). Elite cues and the rapid decline in trust in science agencies on covid-19. *Sociological Perspectives*, 64(5):988–1011.

Hartocollis, A. (2024). The next battle in higher ed may strike at its soul: Scholarship.

Harvard (2024). Francesca gino (harvard business school). https://www.hbs.edu/faculty/Pages/profile.aspx?facId=271812. Accessed: Jan 07 2024.

Hjort, J., Moreira, D., Rao, G., and Santini, J. F. (2021). How research affects policy: Experimental evidence from 2,150 brazilian municipalities. *American Economic Review*, 111(5):1442–80.

Iacobucci, S. and De Cicco, R. (2022). A literature review of bullshit receptivity: Implications for policy makers. 6:23–40.

IEA (2022). Nuclear power and secure energy transitions. https://www.iea.org/reports/nuclear-power-and-secure-energy-transitions. Licence: CC BY 4.0.

Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Medicine*, 2(8).

Kotcher, J. E., Myers, T. A., Vraga, E. K., Stenhouse, N., and Maibach, E. W. (2017). Does engagement in advocacy hurt the credibility of scientists? results from a randomized national survey experiment. *Environmental Communication*, 11(3):415–429.

Krosnick, J. and Alwin, D. (1989). Aging and susceptibility to attitude change. *Journal of Personality and Social Psychology*.

Lee, B. Y. (2020). Nature endorses joe biden, latest top scientific journal to condemn trump's actions.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *The Review of Economic Studies*, 76(3):1071–1102.

Lewis-Kraus, G. (2023). They studied dishonesty. was their work a lie?

Mellström, C. and Johannesson, M. (2008). Crowding out in blood donation: Was titmuss right? *Journal of the European Economic Association*, 6(4):845–863.

Motta, M. (2018). The polarizing effect of the march for science on attitudes toward scientists. *PS: Political Science Politics*, 51:1–6.

NASEM (2016). Genetically engineered crops. *Experiences and Prospects*.

Oliveira, R., Santos, A., and Severnini, E. (2024). Bridging the gap: Mismatch effects and catch-up dynamics under a brazilian college affirmative action program. *Economics of Education Review*, 98:102501.

Pennycook, G., Allan Cheyne, J., Barr, N., Koehler, D. J., and Fugelsang, J. A. (2015). On the reception and detection of pseudo-profound bullshit. *Judgment and Decision Making*, 10(6):549–563.

Rafkin, C., Shreekumar, A., and Vautrey, P.-L. (2021). When guidance changes: Government stances and public beliefs. *Journal of Public Economics*, 196:104319.

Rogger, D. and Somani, R. (2023). Hierarchy and information. *Journal of Public Economics*, 219:104823.

Simohnson, U., Nelson, L., and Simmons, J. (2021a). Evidence of fraud in an influential field experiment about dishonesty.

Simohnson, U., Nelson, L., and Simmons, J. (2021b). Thank you (and an update).

Simohnson, U., Nelson, L., and Simmons, J. (2023a). Data falsificada (part 1): Clusterfake.

Simohnson, U., Nelson, L., and Simmons, J. (2023b). Data falsificada (part 2): My class year is harvard.

Simohnson, U., Nelson, L., and Simmons, J. (2023c). Data falsificada (part 3): The cheaters are out of order.

Simohnson, U., Nelson, L., and Simmons, J. (2023d). Data falsificada (part 4): Forgetting the words.

Sullivan, K. (2020). “i support joe biden’s pro-science agenda”: 81 nobel laureates endorse biden for president — cnn politics.

VoxDev (2024). Affirmative action in brazil’s higher education system. Accessed on Mar 26, 2024.

West, J. D. and Bergstrom, C. T. (2021). Misinformation in and about science. *Proceedings of the National Academy of Sciences*, 118(15):e1912444117.

Zhang, F. J. (2023). Political endorsement by nature and trust in scientific expertise during covid-19. *Nature Human Behaviour*.

Zimmermann, F. (2020). The dynamics of motivated beliefs. *American Economic Review*, 110(2):337–61.