Commitment Requests Do Not Affect Truth-Telling in Laboratory and Online Experiments

Tobias Cagala[a], Ulrich Glogowsky[b,e], Johannes Rincke[c,e], Simeon Schudy[d,e]

[a]*Deutsche Bundesbank, Wilhelm-Epstein-Strasse 14, 60431, Frankfurt am Main, Germany*
[b]*Department of Economics, Johannes Kepler University Linz, Altenbergerstrasse 69, 4040, Linz, Austria*
[c]*Corresponding author. Department of Economics, University of Erlangen-Nuremberg, Lange Gasse 20, 90403, Nuremberg, Germany*
[d]*Department of Economics, Ulm University, Helmholtzstrasse 1, 89081, Ulm, Germany*
[e]*CESifo, Poschingerstrasse 5, 81679, Munich, Germany*

Abstract

Using a standard cheating game, we investigate whether the request to sign a no-cheating declaration affects truth-telling. Our design varies the content of a no-cheating declaration (reference to ethical behavior vs. reference to possible sanctions) and the type of experiment (online vs. offline). Irrespective of the declaration's content, commitment requests do not affect truth-telling, neither in the laboratory nor online. The inefficacy of commitment requests appears robust across different samples and does not depend on psychological measures of reactance.

Keywords: cheating, lying, truth-telling, compliance, commitment, no-cheating rule, no-cheating declaration, commitment request

---

1. Introduction

In many contexts, agents face incentives to misreport private information. Examples include employees overstating their working hours, businesses not disclosing all characteristics of their products, and households and firms understating their income or profit when reporting to the tax authorities. Given that misreporting affects important economic outcomes in all these situations, it is crucial to understand which countermeasures policymakers can use to induce truth-telling.

The standard measure to curb cheating is deterrence, and a large body of empirical literature since Becker (1968) has shown that deterrence works.[1] But how can a principal induce truthful reporting by agents in settings where implementing deterring tools (such as third-party reporting or close monitoring) is too costly or technically impossible? A widely used instrument in such contexts is requesting the agent to commit to certain rules by signing a no-cheating declaration. For example, many universities require students to sign an honor code that spells out the principles of academic integrity.[2] Similarly, when individuals and firms report tax-relevant information, the tax administration commonly requests them to sign a declaration confirming the truthfulness of the submitted information.[3] Furthermore, all Fortune Global 500 corporations have a code of conduct that newly hired staff must sign. These codes frequently include declarations of compliance.[4]

One can think of various channels through which the act of commitment through such declarations could induce more truthful behavior. For example, an intrinsic disutility of cheating shapes many agents' reporting behavior (see, e.g., Gneezy, 2005; Erat and Gneezy, 2012). The commitment may focus the agents' attention on their own moral standards or serve as a reminder. As a result, the perceived disutility of cheating would increase, shifting the tradeoff between truth-telling and misreporting toward more honesty. Likewise, the act of commitment could reduce cheating due to a disutility of breaking a promise (Ellingsen and Johannesson, 2004; Charness and Dufwenberg, 2006). However, there are also reasons to believe that requesting commitment could lead to less truthful reporting. A well-known example of such a channel is psychological reactance. Going back to Brehm

[1]For example, police reduce crime (Levitt, 1997, 2002; Di Tella and Schargrodsky, 2004), auditing and third-party information reporting limit tax evasion (Kleven et al., 2011; Kleven, 2014; Pomeranz, 2015), and close monitoring curbs the mismanagement of publicly funded institutions (Reinikka and Svensson, 2005; Olken, 2007; Ferraz and Finan, 2008, 2011; Bjorkman and Svensson, 2010).

[2]According to the *U.S. News & World Report 2019*, all top 10 U.S. universities have an honor code or code of conduct that explicitly refers to academic integrity. Moreover, four out of the ten require undergraduate students to sign or pledge adherence to this code.

[3]One exemplary country that made use of such a commitment request is Sweden. Before 2002, the Swedish income-tax-return form included the following statement that individuals had to sign: "I promise in honor that the submitted figures are correct and truthful."

[4]For details, see the compliance database of the University of Houston (weblink).

(1966), the theory of reactance states that individuals have a fundamental need for behavioral freedom. This need is activated whenever individuals feel a restriction imposed on their options or actions, leading them to an emotional state characterized by the wish to regain their freedoms through engaging in the restricted activity. In this vein, commitment requests that impose behavioral restrictions could lead individuals to deliberately choose the type of behavior that the request marks as (socially) undesirable.[5] Commitment requests could also weaken the perceived social norm of honesty and thereby decrease the disutility of cheating (Cagala et al., 2023). With competing conceptual frameworks predicting very different effects of commitment requests, it is vital to examine the effects of such a policy.

Against this backdrop, this paper presents pre-registered evidence from controlled economic experiments with almost 700 participants on how requests to sign no-cheating declarations affect misreporting behavior. Our key contribution lies in providing a broad and systematic analysis. The analysis acknowledges that, in practice, policymakers use (a) varying declarations in (b) offline and online contexts. Our paper, hence, not only examines if the decision environment shapes the efficacy of commitment requests but also whether the declarations' contents matter. Specifically, we test how two widely-used no-cheating declarations (one that highlights ethical behavior and one that refers to a no-cheating rule and potential sanctions) affect dishonesty in offline and online decision environments.

Methodologically, we implement a between-subjects design in which participants are asked to sign a no-cheating declaration and later participate in a reporting task in which they can cheat to increase their earnings. Inspired by the idea that effective no-cheating declarations make ethics salient and thereby increase the decision-maker's psychological costs of lying, the first experimental treatment (ETHICS) features a morally-loaded no-cheating declaration. Participants in this treatment sign a declaration to "acknowledge the principles of ethically sound behavior." The second treatment (SANCTION) asks participants to sign a declaration stating that they "will not violate the rules" and also points them to a possible sanction in case of rule violations ("violating the rules can lead to exclusion from future experiments"). While

[5] Researchers established the relevance of reactance in other contexts than commitment requests (see, e.g., the reviews of Miron and Brehm, 2006; Rains, 2013; Steindl et al., 2015).

policymakers rely on such declarations to increase honest behavior, this declaration's effect is ex-ante unclear: On the one hand, the costs of dishonest behavior may increase, and lying may consequently decrease. On the other hand, the declaration conveys a very direct message restricting the behavioral freedom of participants, which may trigger psychological reactance and thereby increase dishonesty (see, e.g., the reviews of Miron and Brehm, 2006; Steindl et al., 2015). We compare these two treatments to a CONTROL condition, in which participants are not asked to sign a no-cheating declaration. We embed all three treatments in simple online and offline cheating games following the idea of Fischbacher and Föllmi-Heusi (2013).[6]

Our main finding is that irrespective of the content of the no-cheating declaration (reference to ethical behavior vs. possible sanction) and the type of experiment (online vs. offline sessions), requesting participants to sign a no-cheating declaration has no discernable effect on truth-telling. Across experimental conditions and in both environments (laboratory and online), the share of participants who cheat is close to the overall mean of 32 percent. Further, the treatment effect sizes do not systematically vary with participants' psychological reactance.

Literature:. Our paper contributes to (a) the literature on commitment requests and (b) the laboratory-experimental literature on oaths and moral reminders. Most closely related, Cagala et al. (2023) study the effects of a no-cheating declaration in the context of academic exams, where the resulting punishment for cheating is clear, and expectations about others' honesty are high.[7] The paper finds that a no-cheating declaration does not result in less cheating but may even backfire as it can shift students' expectations about their peers' honesty.[8] Our findings complement this work by highlighting

[6]The online experiment included an additional treatment condition with a neutrally framed no-cheating declaration (see Section 2.2 for details). Misreporting in this condition does not differ from misreporting in the control condition. To increase statistical power, we decided against implementing this treatment in the lab and used all the available observations for the other two treatments.

[7]Their declaration read as follows: "I hereby declare that I will not use unauthorized materials during the exam. Furthermore, I declare neither to use unauthorized aid from other participants nor to give unauthorized aid to other participants."

[8]In a similar vein, the Behavioural Insights Team (2012) finds suggestive evidence that moving a no-cheating declaration from the bottom to the top of a form to apply for a tax discount may increase rather than decrease fraud. Koretke (2017) examines in a small-

that declarations relating to rule violations and sanctions are not increasing honesty in environments with less clear punishment rules for misbehavior and lower priors about others' honesty. Further, we extend the literature by documenting that the inefficacy prevails across declarations and contexts (online vs. offline).

As mentioned, there is also laboratory-experimental literature that mainly focused on the impacts of oaths and moral reminders on honesty. It presents mixed results. Our contribution to this literature is to focus on commitment requests and to provide a systematic analysis of such requests by varying (a) the declaration and (b) the decision environment. One of the previous studies is, for example, Beck et al. (2020), who contrast (among other treatments) a baseline condition of the Fischbacher and Föllmi-Heusi (2013) task (n=39) with a moral-awareness treatment. In this treatment, participants confirm with their signature that the data they provide regarding their actions during the experiment align with the principle of honesty and that they do not lie to enrich themselves (n=29). Using substantially smaller samples than our study, they find, on average, lower reports in the latter condition ($p < 0.05$). Jacquemet et al. (2018) study the efficacy of truth-telling oaths on honesty in a sender-receiver game (n=60 in each treatment condition).[9] Akin to our results, they find that having subjects sign a truth-telling oath before participating in a neutrally framed lying game leaves truth-telling behavior unchanged. Simultaneously to and independently of our work, Schild et al. (2019) implemented a large-scale online study on moral reminders, visibility, and self-engagement.[10] They compare promise and no-promise conditions and do not find statistically significant differences in the probability of dishonesty if misbehavior is observable.[11] Complementing and extending this work, we show that not only commitment requests related to ethical behavior but also declarations related to rule violations and potential punishment

scale study ($n = 48$) if the type of commitment matters (verbal vs. written) and finds no effects.

[9]In their study, participants voluntarily sign a form asking them "to swear upon [their] honor that, during the whole experiment, [they] will tell the truth and always provide honest answers."

[10]They manipulated self-engagement by asking participants to "promise that the information [they] are providing is true" before a cheating task.

[11]In a condition where cheating is unobserved, the study finds a statistically significant reduction of 8 percentage points in the (estimated) probability of dishonesty due to self-engagement.

have no effects on reporting behavior. Moreover, we not only focus on online but also on offline contexts.

Finally, given that two influential studies on how moral reminders and commitment requests affect cheating were recently found to suffer from issues of replicability and research integrity, our systematic analysis of the effects of commitment requests is all the more warranted. In the first of these studies, Mazar et al. (2008) test the effects of a moral-reminder treatment that nudged participants to recall the Ten Commandments (Mazar et al., 2008, Experiment 1). Verschuere et al. (2018) fail to confirm that such a treatment can increase honesty in a large-scale replication exercise. In the second study, Shu et al. (2012) report evidence suggesting that principals can increase honesty by asking agents to sign a no-cheating declaration *before* rather than *after* providing information. However, these results did not replicate (Kristal et al., 2020), and Simonsohn et al. (2021) provided evidence that questioned the data's validity and ultimately led to the original study's retraction.

The structure of the paper is as follows. Section 2 describes the experimental design, Section 3 discusses the results, and Section 4 concludes. The Online Appendix provides theoretical considerations and supplementary results.

2. Experimental Design

This subsection introduces the designs of our laboratory (Subsection 2.1) and online experiments (Subsection 2.2). We conducted the lab experiment in 2022 and the online experiment in 2020.

2.1. Laboratory Experiment

We first describe the basics of our experimental design and then the treatments.[12] The design consists of two parts: a survey and a cheating game. The

[12]Prompted by referee comments on an earlier version of this paper concerning experimenter-demand effects and statistical power, we re-worked the experimental design and pre-registered the new experiment under https://doi.org/10.1257/rct.6700. Because of the COVID-19 pandemic, laboratories were closed back then, and we preregistered the data collection as an online experiment. After the re-opening of laboratories in Germany, we pre-registered the design and the data analysis of the laboratory experiment under https://doi.org/10.1257/rct.9683.

cheating game follows the computerized experiment of Abeler et al. (2019).[13]

Part I: Survey:. After participants entered the laboratory, the experimenter informed them that the session consisted of two parts. In the first part, participants received a payoff of €4 for answering a 15-minute survey on the German inheritance tax schedule (see Appendix A.1 for details). We added this part to the experiment for two reasons. First, by placing other elements before the cheating decision, we followed the standard experimental protocol in the literature (see, e.g, Fischbacher and Föllmi-Heusi, 2013; Kajackaite and Gneezy, 2017). Second, and more importantly, we included this survey to introduce our commitment requests more naturally and mitigate experimenter demand effects. In the respective sessions, the experimenter placed the printed declaration at the participants' workplaces before they entered the laboratory and reminded them to sign the "declaration concerning the behavioral rules in the laboratory" right at the beginning of the session. Therefore, we connected the commitment request to the entire session rather than to the cheating experiment.

Part II: Cheating Game:. At the beginning of the session's second part, the participants read instructions on the computer screen (see Appendix A.2). The instructions informed participants that the experiment would start with a computerized random draw of a number between one and six that they would have to self-report. Participants also learned from the instructions that their additional payoff (i.e., the payoff in addition to the fixed payment for participating in the survey) would be €5 if they reported a five and zero if they reported a number from the set $\{0, 1, 2, 3, 4, 6\}$.

The computerized random draw simulated the drawing a chip from an envelope. Participants first saw an envelope containing six chips numbered between one and six on their screen (see Appendix A.3 for screenshots). They then clicked a button to start the draw. The chips were shuffled for a few seconds, and one randomly selected chip fell out of the envelope. On the next screen, participants were asked to report their draw by entering the number into a field on the screen.[14] After the participants had reported their

[13] We thank the authors for providing code to replicate the computerized draw in their experiment. We programmed the experiment with z-Tree (Fischbacher, 2007) and recruited participants with ORSEE (Greiner, 2015).

[14] Before reporting their draw, participants could also click a button to show the instruc-

numbers, the experimenter called them by their computer number and paid them anonymously for both parts of the session.

The fact that we computerized the random draw makes cheating observable at an individual level. This design element comes with a much higher statistical power than approaches that identify cheating by evaluating the empirical distribution of self-reports against the expected distribution under truthful reporting (see, e.g., Fischbacher and Föllmi-Heusi, 2013). Akin to other studies that have used reporting tasks with observable decisions involving dishonesty (see, e.g., Gneezy et al., 2018; Kocher et al., 2018), our instructions did not explicitly state that the experimenter could observe misreporting at the individual level. Nevertheless, we consider it reasonable that participants (a) were aware that the experimenter could observe cheating at the individual level but (b) still did not expect punishment for misreporting.[15] The latter is because our instructions highlighted that a subject's payoff depended exclusively on their report and, thus, clarified that misreporting does not lead to any (immediate) monetary sanctions.[16]

Treatments:. We implemented a CONTROL and two treatment conditions. Participants in the CONTROL condition did not sign a no-cheating declaration. By contrast, participants in the treatments signed such a declaration right after they entered the laboratory and took their seats. The paper with the declaration displayed a short preamble highlighting that experiments at

tions and the payoff structure again. They could also click a button to display the result of the random draw again.

[15]This observation is also reflected by the fact that in settings with observable cheating decisions, few participants tend to lie partially conditional on lying (see Gneezy et al., 2018; Kocher et al., 2018).

[16]This procedure and the fraction of cheaters observed in our experiment (32 percent on average) are similar to other studies using reporting paradigms in observed environments. For instance, in Gneezy et al. (2018, p. 440), 26–33 percent of participants misreported the observed outcome, and Kocher et al. (2018, p. 4000) found that 31–41 percent of individuals did not report truthfully. Nevertheless, we cannot preclude that (some) participants disliked that the instructions were not explicitly stating that the random draw was recorded. If so, this may have motivated them to reciprocate negatively by misreporting. However, the share of cheaters in our data is much lower than what we typically observe in experimental designs that do not record random draws. For instance, in Fischbacher and Föllmi-Heusi (2013), only 39 percent of participants are fully honest (compared to 68 percent in our setting). Hence, we consider it unlikely that our setting triggered misreporting due to negative reciprocity.

the respective laboratory are subject to certain behavioral standards and/or rules. Below the preamble, the paper included a brief declaration. The treatments varied the declarations' content. In our first treatment, the ETHICS condition, the declaration that followed the preamble[17] read:

> "*I hereby acknowledge the principles of ethically sound behavior.*"

The treatment aimed at making the ethical dimension of cheating salient without communicating behavioral restrictions that could trigger reactance. By contrast, in our second treatment, called SANCTION, the declaration that followed the preamble[18] read:

> "*I hereby declare that I will not violate the rules described in the instructions. Violating the rules can lead to exclusion from future experiments.*"

As discussed in the introduction and clarified in Online Appendix B (conceptual framework), the effect of such a no-cheating declaration is unclear. On the one hand, highlighting a potential sanction in the case of non-compliance could trigger reactance and, thereby, lead to more cheating. On the other hand, the declaration could also increase honest behavior, either by acting as a moral reminder or because the threat of a sanction may lead subjects to update their beliefs about sanctions for false reports.[19]

Online Survey to Elicit Reactance:. To test the heterogeneity of possible treatment effects with respect to the participants' degree of reactance as a trait, we elicited Hong's Psychological Reactance Scale (Hong, 1992; De las

[17]The preamble read: "The [name of laboratory] adheres to the ethical standards that were defined, e.g., by the German Research Foundation. One of the principles of ethically sound behavior is that data and findings must not be falsified. Today's experiment is subject to the stated standards."

[18]The preamble read: "At the [name of laboratory], participants participating in experiments have to adhere to certain rules. One of the rules requires participants to follow the behavioral guidelines provided in the instructions for the experiment. Please sign the following declaration referring to this rule."

[19]As discussed, the instructions clearly stated that participants would receive the additional payoff if reporting a five. Hence, the instructions did not entail any signal that giving a false report would trigger monetary punishments. However, we cannot fully preclude that the SANCTION treatment triggered a higher perceived risk of exclusion from future experiments in case of a false report.

Cuevas et al., 2014). The scale consists of 14 statements that approximate the degree to which one person shows reactance (5-point Likert scale). Online Appendix C provides the complete list of statements. We collected the reactance data two weeks *after* the laboratory experiment using an additional online survey. After receiving an email invitation, participants had 48 hours to answer the questionnaire. Answering the online survey took about five minutes, and participants received a fixed payoff of €2. To obfuscate the purpose of the online survey, we mixed the reactance questions with 15 questions commonly used to elicit the Big 5 personality traits (Gerlitz and Schupp, 2005).

2.2. Online Experiment

Design and Treatments: . All core aspects of the experimental design of the online experiment were identical to the one of the laboratory experiment. We, again, implemented a design with the same two parts (survey and cheating game). Also, the treatments were the same, and we, again, invited participants from a subject pool typically used to recruit subjects for laboratory experiments. Only a few aspects of our design were different compared to the laboratory: First, naturally, we had to present the declarations in the ETHICS and SANCTION conditions on the computer screen rather than on paper. Participants signed the declaration (before Part I started) by typing their first and second names into a text field. Second, we conducted the reactance survey two weeks before the online experiment (rather than two weeks after the laboratory experiment).[20] Third, the online experiment included a third treatment group with a neutrally framed no-cheating declaration.[21] In line with our main result that commitment requests are ineffective, this treatment also did not affect cheating (see Online Appendix A). When designing our laboratory experiment in 2022, the combined subject pools (of

[20]We changed the ordering in the laboratory experiment (relative to the online experiment) to ensure (by design) that survey responses cannot affect behavior in the cheating game. Further, to compensate participants in the laboratory experiment appropriately and to comply with the local laboratory rules, we had to increase the flat payment for Part 1 in the laboratory experiment by 1 Euro.

[21]See the AEA RCT registry entry at `https://doi.org/10.1257/rct.6700` for details. The declaration read: "I hereby declare that I will not violate the rules described in the instructions." Online Appendix Table A.1, Column 7, in Online Appendix A reports results for this additional treatment in the online experiment.

the three used laboratories) were not large enough to implement three treatments with sufficient statistical power. We, hence, decided to focus on the two treatments with the largest expected effects (ETHICS and SANCTION).

2.3. Further Details

Between July and November 2022, we conducted our laboratory experiments in two different laboratories: the MELESSA laboratory at the University of Munich (361 observations) and the Lakelab laboratory at the University of Konstanz (149 observations). The sessions lasted about 45 minutes, including time for the participants' payment. Participants completing all parts of the experiment (including the online questionnaire) received an average payoff of €14.1 (including a show-up fee of €6 for coming to the laboratory). In December 2020, we collected 323 observations for our online experiment using the subject pool of the LERN laboratory at the University of Erlangen-Nuremberg. These participants earned, on average, €13.3, including a €6 show-up fee for participating in the online session. The overall number of observations is thus 833.

3. Results

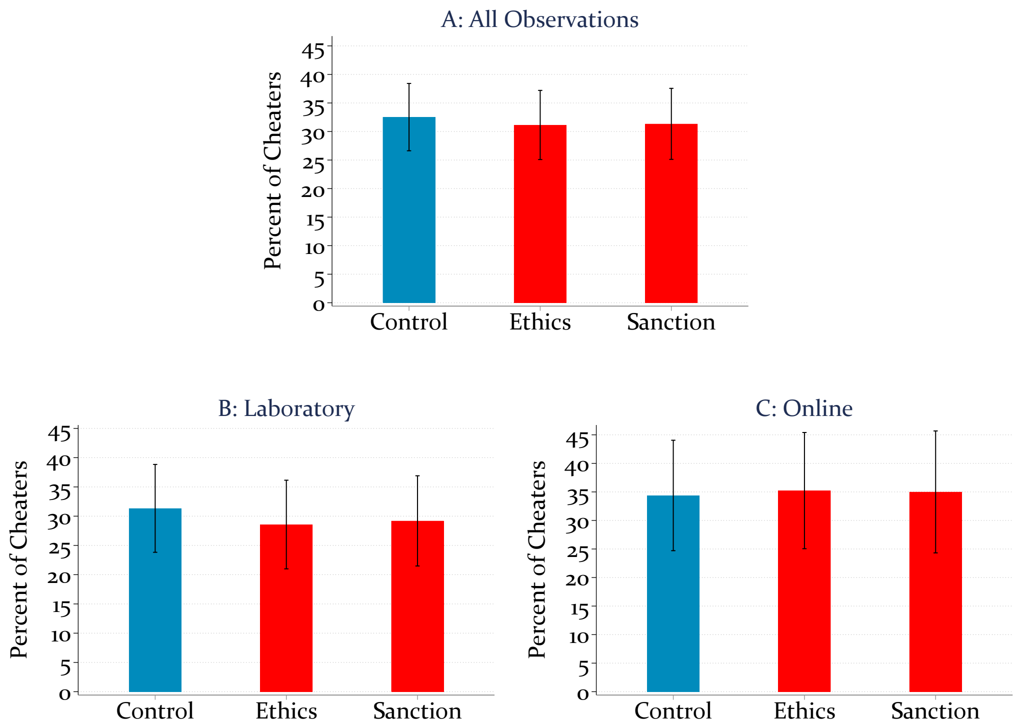
Main Result:. Figure 1 displays our main finding: The requests to sign no-cheating declarations in online and laboratory experiments do not affect cheating. To demonstrate this insight, we analyze behavior in the sample of all participants who did not draw a five in the random draw and, thus, had a profitable option to cheat. This leaves us with an effective sample size of 691 out of the 833 observations we collected.[22] From these 691 participants, 246 received the CONTROL group, 228 in the ETHICS treatment, and 217 in the SANCTION treatment.[23]

To demonstrate our main finding, Panel A of Figure 1 pools the data from the laboratory and online experiments ($N = 691$) and shows the share of cheaters across the treatment conditions. This share was 31.1 percent in the ETHICS and 31.3 percent in SANCTION conditions, slightly less than the share of 32.5 percent in the CONTROL condition. These minor differences

[22]The effective sample sizes by subject pools are 299 at MELESSA, 128 at Lakelab, and 264 at LERN (online).

[23]Due to differences in participation rates in the sessions assigned to each treatment and the nature of the random draw, the treatment groups differ in size.

Figure 1: Cheating Behavior by Treatment



Notes: This figure shows the share of individuals (in percent) who cheated in each experimental condition. Panel A displays the results for the pooled sample (pooling over all participants in the online and laboratory experiments; $N = 691$). Sample size by treatment group: 246 in CONTROL, 228 in ETHICS, and 217 in SANCTION. Panel B focuses on the laboratory experiment ($N = 427$) and Panel C on the online experiment ($N = 264$). The spikes indicate 95% confidence intervals of unconditional means.

are neither economically nor statistically significant. Using non-parametric χ^2-tests comparing the proportion of cheaters across treatments, the p-values amount to $p = 0.747$ (CONTROL vs. ETHICS), $p = 0.785$ (CONTROL vs. SANCTION), and $p = 0.964$ (ETHICS vs. SANCTION). Linear probability models with robust standard errors yield similar results (all p-values > 0.74; see Online Appendix Table A.1 in Online Appendix A).

The second row of Figure 1 repeats the same analysis separately for the laboratory experiment (Panel B) and the online experiment (Panel C). Both panels show very similar results. Most importantly, none of the differences between the experimental conditions are statistically significant. Considering only the data from our laboratory experiment, the p-values of χ^2-tests are $p = 0.608$ (CONTROL vs. ETHICS), $p = 0.694$ (CONTROL vs. SANCTION), and $p = 0.909$ (ETHICS vs. SANCTION). The respective values for our online experiments are $p = 0.903$ (CONTROL vs. ETHICS), $p = 0.931$ (CONTROL vs. SANCTION), and $p = 0.975$ (ETHICS vs. SANCTION).

Further Analyses:. In the Online Appendix, we provide two additional analyses. First, we test the robustness of the findings obtained from Figure 1 by employing regressions. All our results are confirmed (Online Appendix Table A.1). Second, following our preregistration,[24] we also shed light on potential treatment heterogeneity based on measures of participants' psychological reactance. Particularly, in Online Appendix B, we integrate the concept of psychological reactance into the simple conceptual framework for cheating and lying behavior by Kajackaite and Gneezy (2017). The altered framework predicts that commitment requests that impose behavioral restrictions could lead reactant individuals to deliberately choose the type of behavior that the request marks as undesirable (e.g., restricting the freedom of choice by a rule with a potential sanction may lead to more rule violations). We then test this hypothesis by classifying individuals according to their reactance type (Hong, 1992) and studying whether less and more reactant types respond differently to the treatment. Our analyses indicate that the treatment effects do not systematically vary across the participants' types (Online Appendix Table B.1). In particular, we do not find support for the hypothesis that the SANCTION treatment triggers more cheating by more reactant types (nor any other systematic treatment heterogeneity).

4. Conclusion

Universities, firms, and public institutions frequently require individuals to commit to truthful reporting of private information. A common implementation of such commitment requests is to let individuals sign a no-cheating declaration. However, from a theoretical perspective it is unclear how such requests affect individual behavior. This paper implements a laboratory and an online experiment with almost 700 participants to test empirically how two types of no-cheating declarations affect truth-telling when participants face a profitable option to misreport private information. The requests we

[24] https://doi.org/10.1257/rct.9683.

study build on two commonly used strategies to alter the (psychological) costs of lying: (a) alluding to a principle of ethically sound behavior and (b) highlighting common rules and possible sanctions. Our main finding is that irrespective of the content of the no-cheating declaration (reference to ethical behavior vs. possible sanction) and the type of experiment (online vs. offline sessions), requesting participants to sign a no-cheating declaration has no discernable effect on truth-telling.

Importantly, our results are robust across three different subject pools and two types of decision environments (online and offline). Moreover, the (average) inefficacy of commitment requests is not a result of heterogenous reactions due to individuals' psychological reactance. Complementing and extending previous work, we, thus, show that commitment requests highlighting ethical behavior or potential sanctions are unlikely to improve outcomes in settings such as ours. Recent work on oaths and moral reminders (Jacquemet et al., 2018; Schild et al., 2019), however, suggests that aspects of the decision environment may matter (e.g., whether a lie is framed as a lie or if lying is observable). Future studies should, hence, systematically analyze whether these and other environmental features may render commitment requests effective.

5. Acknowledgments

We are thankful for financial support by the Deutsche Forschungsgemeinschaft (DFG) through CRC/TRR190 (project number: 280092119). Ulrich Glogowsky also thanks Alan Auerbach and the DFG for supporting and funding his research stay at University of California (Berkely), during which this paper was partly written. We appreciate helpful comments from two anonymous referees and seminar participants at various places. The paper represents the authors' personal opinions and does not necessarily reflect the views of the Deutsche Bundesbank or its staff. Declarations of interest: none. The data and code needed to replicate our analysis are available at https://doi.org/10.7910/DVN/8DKFYO.

References

Abeler, J., Nosenzo, D., Raymond, C., 2019. Preferences for Truth-Telling. Econometrica 87, 1115–1153.

Beck, T., Bühren, C., Frank, B., Khachatryan, E., 2020. Can Honesty Oaths, Peer Interaction, or Monitoring Mitigate Lying? Journal of Business Ethics 163, 467–484.

Becker, G.S., 1968. Crime and Punishment: An Economic Approach. Journal of Political Economy 76, 169–217.

Behavioural Insights Team, 2012. Applying Behavioural Insights to Reduce Fraud, Error and Debt. Technical Report. Cabinet Office, London.

Bjorkman, M., Svensson, J., 2010. When Is Community-Based Monitoring Effective? Evidence from a Randomized Experiment in Primary Health in Uganda. Journal of the European Economic Association 8, 571–581.

Brehm, J.W., 1966. A Theory of Psychological Reactance. Oxford: Academic Press.

Cagala, T., Glogowsky, U., Rincke, J., 2023. Detecting and Preventing Cheating in Exams: Evidence from a Field Experiment. *Journal of Human Resources*, forthcoming.

Charness, G., Dufwenberg, M., 2006. Promises and Partnership. Econometrica 74, 1579–1601.

De las Cuevas, C., Penate, W., Betancort, M., de Rivera, L., 2014. Psychological Reactance in Psychiatric Patients: Examining the Dimensionality and Correlates of the Hong Psychological Reactance Scale in a Large Clinical Sample. Personality and Individual Differences 70, 85–91.

Di Tella, R., Schargrodsky, E., 2004. Do Police Reduce Crime? Estimates Using the Allocation of Police Forces After a Terrorist Attack. American Economic Review 94, 115–133.

Ellingsen, T., Johannesson, M., 2004. Promises, Threats and Fairness. Economic Journal 114, 397–420.

Erat, S., Gneezy, U., 2012. White Lies. Management Science 58, 723–733.

Ferraz, C., Finan, F., 2008. Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes. Quarterly Journal of Economics 123, 703–745.

Ferraz, C., Finan, F., 2011. Electoral Accountability and Corruption: Evidence from the Audits of Local Governments. American Economic Review 101, 1274–1311.

Fischbacher, U., 2007. z-Tree: Zurich Toolbox for Ready-made Economic Experiments. Experimental Economics 10, 171–178.

Fischbacher, U., Föllmi-Heusi, F., 2013. Lies In Disguise–An Experimental Study On Cheating. Journal of the European Economic Association 11, 525–547.

Gerlitz, J.Y., Schupp, J., 2005. Zur Erhebung der Big-Five-Basierten Persönlichkeitsmerkmale im SOEP. DIW Research Notes, 4.

Gneezy, U., 2005. Deception: The Role of Consequences. American Economic Review 95, 384–394.

Gneezy, U., Kajackaite, A., Sobel, J., 2018. Lying Aversion and the Size of the Lie. American Economic Review 108, 419–53.

Greiner, B., 2015. Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE. Journal of the Economic Science Association 1, 114–125.

Hong, S.M., 1992. Hong's Psychological Reactance Scale: A further Factor Analytic Validation. Psychological Reports 70, 512–514.

Jacquemet, N., Luchini, S., Rosaz, J., Shogren, J.F., 2018. Truth-Telling under Oath. Managament Science 65, 426–438.

Kajackaite, A., Gneezy, U., 2017. Incentives and Cheating. Games and Economic Behavior 102, 433–444.

Kleven, H.J., 2014. How Can Scandinavians Tax So Much? Journal of Economic Perspectives 28, 77–98.

Kleven, H.J., Knudsen, M.B., Kreiner, C.T., Pedersen, S., Saez, E., 2011. Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark. Econometrica 79, 651–692.

Kocher, M.G., Schudy, S., Spantig, L., 2018. I Lie? We Lie! Why? Experimental Evidence on a Dishonesty Shift in Groups. Management Science 64, 3995–4008.

Koretke, S., 2017. The Effect of Verbal versus Written Reminder on Cheating Mimeo.

Kristal, A.S., Whillans, A.V., Bazerman, M.H., Gino, F., Shu, L.L., Mazar, N., Ariely, D., 2020. Signing at the Beginning Versus at the End Does Not Decrease Dishonesty. Proceedings of the National Academy of Sciences 117, 7103–7107.

Levitt, S.D., 1997. Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime. American Economic Review 87, 270–290.

Levitt, S.D., 2002. Using Electoral Cycles in Police Hiring to Estimate the Effects of Police on Crime: Reply. American Economic Review 92, 1244–1250.

Mazar, N., Amir, O., Ariely, D., 2008. The Dishonesty of Honest People: A Theory of Self-Concept Maintenance. Journal of Marketing Research 45, 633–644.

Miron, A.M., Brehm, J.W., 2006. Reactance Theory – 40 Years Later. Zeitschrift fuer Sozialpsychologie 37, 9–18.

Olken, B.A., 2007. Monitoring Corruption: Evidence from a Field Experiment in Indonesia. Journal of Political Economy 115, 200–249.

Pomeranz, D., 2015. No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax. American Economic Review 105, 2539–2569.

Rains, S.A., 2013. The Nature of Psychological Reactance Revisited: A Meta-Analytic Review. Human Communication Research 39, 47–73.

Reinikka, R., Svensson, J., 2005. Fighting Corruption to Improve Schooling: Evidence from a Newspaper Campaign in Uganda. Journal of the European Economic Association 3, 259–267.

Schild, C., Heck, D.W., Ścigała, K.A., Zettler, I., 2019. Revisiting REVISE: (Re)Testing Unique and Combined Effects of REminding, VIsibility, and SElf-Engagement Manipulations on Cheating Behavior. Journal of Economic Psychology 75, 102161.

Shu, L.L., Mazar, N., Gino, F., Bazerman, M.H., Ariely, D., 2012. Signing at the Beginning Makes Ethics Salient and Decreases Dishonest Self-Reports in Comparison to Signing at the End. Proceedings of the National Academy of Sciences 109, 15197–15200.

Simonsohn, U., Simmons, J., Nelson, L., 2021. Evidence of Fraud in an Influential Field Experiment About Dishonesty. Data Colada, http://datacolada.org/98.

Steindl, C., Jonas, E., Sittenthaler, S., Traut-Mattausch, E., Greenberg, J., 2015. Understanding Psychological Reactance: New Developments and Findings. Zeitschrift fuer Psychologie 223, 205–214.

Verschuere, B., Meijer, E.H., Jim, A., Hoogesteyn, K., Orthey, R., McCarthy, R.J., Skowronski, J.J., 2018. Registered Replication Report on Mazar, Amir, and Ariely (2008). Advances in Methods and Practices in Psychological Science 1, 299–317.

Appendix A. Experimental Instructions and Screenshots

Appendix A.1. Instructions: Survey

Figure A.1: **Survey: Instructions, Questions, and Solutions**

Questionnaire 1

Thank you for participating in today's session!

General information

Today's study consists of a laboratory session and an *online survey*. For completing the *online survey* within 14 days, you will receive 2 euros. For participating in today's session, you will receive 6 euros. We will transfer the 6+2 euros to your bank account after the completion of the study.

The session today consists of two *independent* parts. In the first part, you will answer a *questionnaire*. For this, you will receive 4 euros. The second part is an *experiment* where you can earn additional money. After today's session, you will receive your payment for both parts in cash.

Please bring all written documents with you when you go to payout!

First Part: Questionnaire

In the following, we kindly ask you to answer nine questions regarding the calculation of the German inheritance tax. You have 15 minutes to complete the questionnaire. You may answer the questions by means of an internet search. A calculator will also be provided.

The questions concern the calculation of inheritance tax for the following case:

In the year 2021, Claudia inherits private assets from her deceased husband (i.e. she receives neither business assets nor real estate). According to the law, she is entitled to a personal allowance of €500,000 and a special allowance of €256,000.

1. Assume that Claudia inherits private assets worth **€806,000**. The taxable inheritance (assets minus allowances) is therefore 806,000 - 500,000 - 256,000 = **€50,000**. Calculate the inheritance tax payment in this case. Provide your result below.

   Your answer: ____________________

2. By how much does the calculated tax payment increase (relative to the situation in 1.) if the inheritance rises by **€100 from €806,000 to €806,100** (taxable inheritance: €50,100)?

   Your answer: ____________________

3. Again, consider the increase of **€100** in 2. What percentage of these additional **100** euros must be paid as taxes? Answer this question using your answer to 2.

   Your answer: ____________________

Figure A.1: **Survey: Instructions, Questions, and Solutions**

4. Now, assume Claudia inherits private assets worth **€836,000**. The taxable inheritance is thus **€80,000**. Calculate the inheritance tax payment in this case. Take the *hardship compensation* into account. Provide your result below.

   Your answer: ____________________

5. By how much does the calculated tax payment increase (relative to the situation in 4.) if the inheritance rises by **€100** from **€836,000** to **€836,100** (taxable inheritance: **€80,100**)? Once again, take the *hardship compensation* into account.

   Your answer: ____________________

6. Again, consider the increase of **€100** in 5. What percentage of these additional **100** euros must be paid as taxes? Answer this question using your answer to 5.

   Your answer: ____________________

7. Now, assume Claudia inherits private assets worth **€850,000**. The taxable inheritance is thus **€94,000**. Calculate the inheritance tax payment in this case. Provide your result below.

   Your answer: ____________________

8. By how much does the calculated tax payment increase (relative to the situation in 7.) if the inheritance rises by **€100** from **€850,000** to **€850,100** (taxable inheritance: **€94,100**)?

   Your answer: ____________________

9. Again, consider the increase of **€100** in 8. What percentage of these additional **100** euros must be paid as taxes? Answer this question using your answer to 8.

   Your answer: ____________________

10. How many of the previous 9 questions do you think you answered correctly? Check the respective box.

☐ 1	☐ 4	☐ 7
☐ 2	☐ 5	☐ 8
☐ 3	☐ 6	☐ 9

11. Age: ____________

12. Gender: ____________

13. Nationality: ____________

14. Field of study: ____________

15. Computer number: ____________

Appendix A.2. Instructions: Cheating Game

Instructions

Thank you for participating in today's experiment!

Please read the instructions carefully. For answering the questionnaire, you will receive **4 Euro** (first part of today's session). There is a possibility to earn another **5 Euro** in the following experiment (second part of today's session).

For showing up today, you will additionally receive **6 Euro** and for completing the online survey, you will receive **2 Euro**. We will transfer the **6+2 Euro** for your participation (in total **8 Euro**) after you have completed all parts of the experiment.

For the first and the second part of today's experiment, you will receive your payoffs in cash (after the completion of the second part of this session). Also note that this is a computer-based experiment. The data will be analyzed anonymously.

Continue

Instructions

Please read the instructions now. When you have finished, click the CONTINUE button.

You will then see **six chips** with the numbers **1, 2, 3, 4, 5, and 6**. Click the START button. The chips will be placed in the envelope. The envelope will be shuffled a couple of times. Then one of the chips will be drawn randomly, and this particular chip will fall out of the envelope.

Please enter the number you have drawn into the field provided for this purpose. You will receive **0 Euro** if you enter the numbers **1, 2, 3, 4, or 6**. You will receive **5 Euro** if you enter a **5**.

Your payment will be determined as follows:

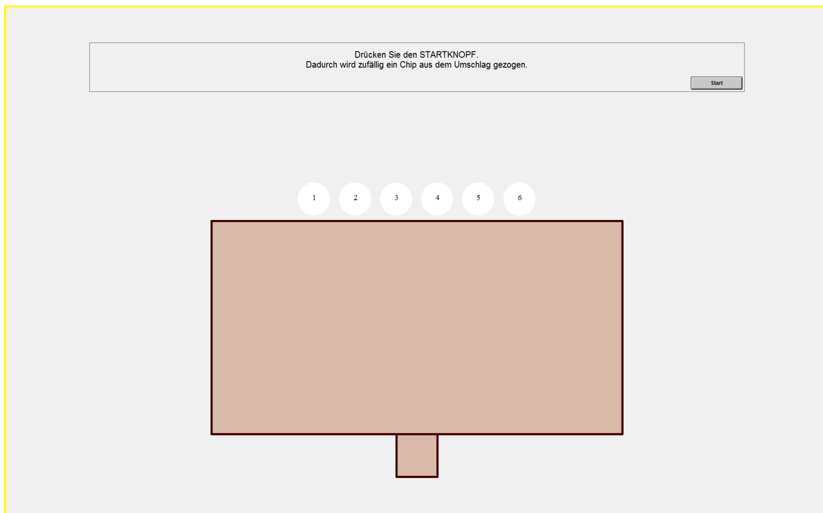
Number	1	2	3	4	5	6
Payoff	0 Euro	0 Euro	0 Euro	0 Euro	5 Euro	0 Euro

Once you have entered your number, you will be called to receive your payment.
Please hand over all written material once you leave the laboratory.

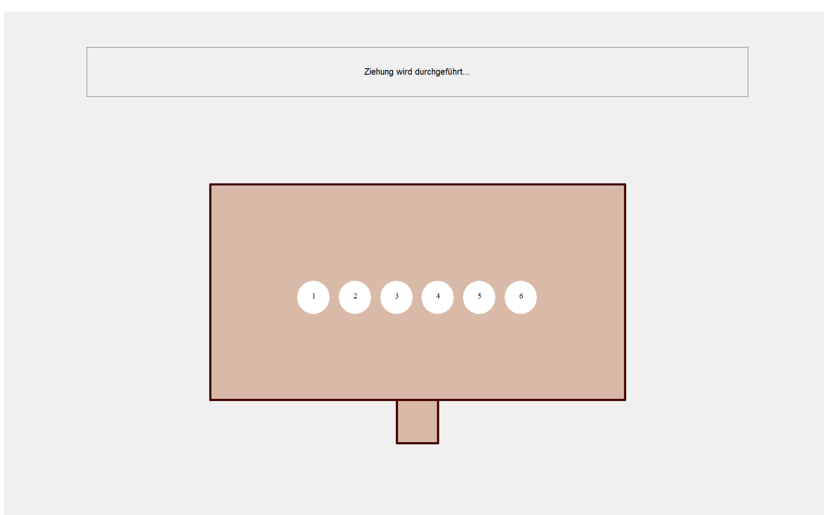
Continue

Appendix A.3. Instructions: Random Draw

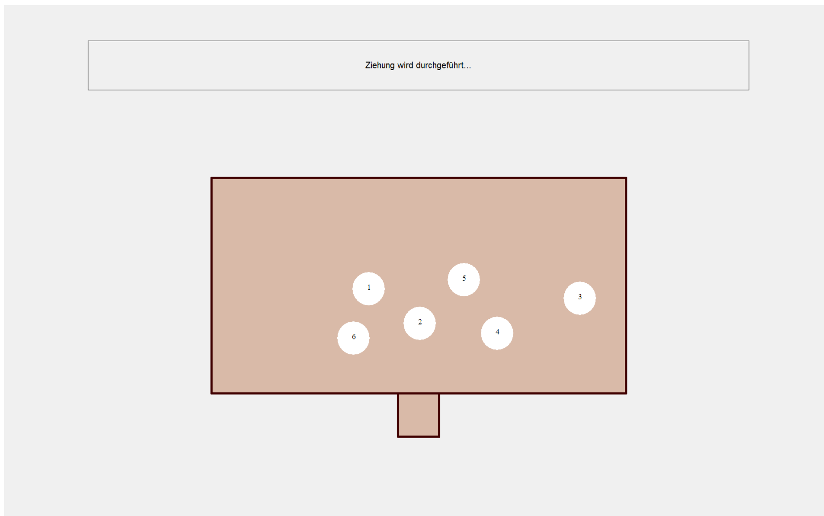
Six chips above envelope



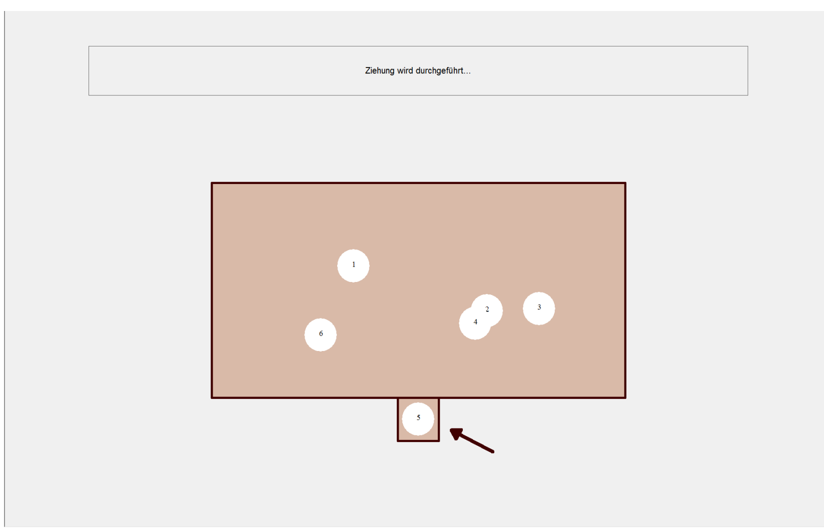
Six chips in envelope



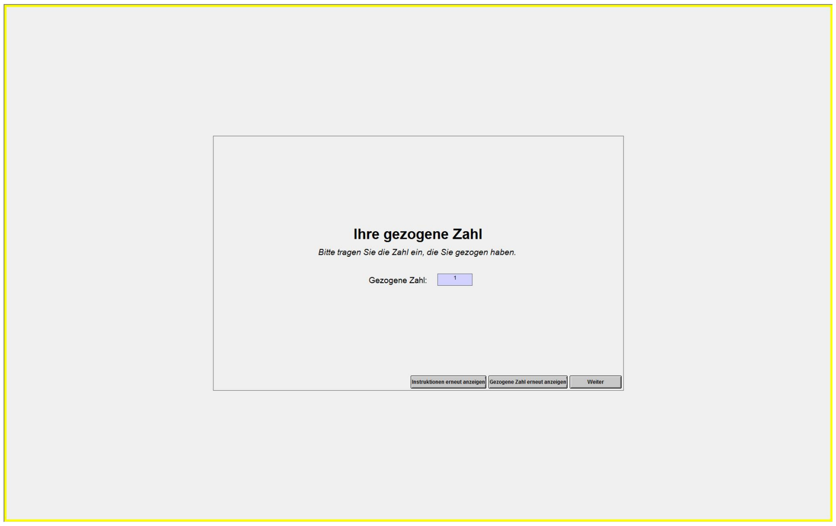
Chips are shuffled



One chip falls out of the envelope



Participants report number



Online Appendix A. Regression Analyses

Table A.1: Treatment Effects on Cheating: Regression Results

	All Observations			Laboratory		Online	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ethics Treatment	-0.014	-0.014	-0.017	-0.028	-0.033	0.009	0.009
	(0.043)	(0.043)	(0.043)	(0.054)	(0.054)	(0.071)	(0.071)
Sanction Treatment	-0.012	-0.011	-0.011	-0.021	-0.022	0.006	0.006
	(0.043)	(0.043)	(0.043)	(0.054)	(0.054)	(0.072)	(0.072)
Neutral Treatment							0.016
							(0.071)
Online		0.051	0.027				
		(0.037)	(0.040)				
Lab Konstanz			-0.080*		-0.081*		
			(0.047)		(0.047)		
N. of obs.	691	691	691	427	427	264	353
Mean control group	0.325	0.325	0.325	0.313	0.313	0.344	0.344

Notes: This table shows linear probability models using as dependent variable an indicator for participants who cheated. Columns (1) to (3) use all observations on participants who had a profitable option to cheat. The omitted category for subject pools in Column (3) is Lab Munich. Columns (4) and (5) use only observations collected in laboratory sessions (Konstanz and Munich). Columns (6) and (7) use only observations collected online (subject pool of laboratory Nuremberg). Column (7) also reports the effect of the additional treatment ("Neutral") that we implemented only in the online experiment but not in the laboratory experiment. This treatment used a neutrally framed commitment request. For further details, see the AEA RCT registry entry at https://doi.org/10.1257/rct.6700 and https://doi.org/10.1257/rct.9683. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Online Appendix B. Psychological Reactance and Heterogeneous Treatment Effects

Following our preregistration[25], we also shed light on potential treatment heterogeneity regarding measures of participants' psychological reactance. Below, we first propose a simple conceptual framework for cheating and lying behavior. This framework is an extension of Kajackaite and Gneezy (2017) and nests possible explanations for negative and positive effects of commitment requests. Second, we analyze empirically, whether heterogeneous treatment effects are observed, using measures of psychological reactance by Hong (1992).

Suppose an agent faces a binary decision to either cheat or not. She observes the state of nature t and then self-reports the state. The agent has two option. She can either report the true state t or report a false state t'. The monetary payoff from stating t is m_t and from reporting t' is $m_{t'}$. This results in a monetary benefit of cheating of $m_{t'} - m_t > 0$. With $p(m_{t'}, m_t)$ denoting the perceived probability of punishment and $s(m_{t'}, m_t)$ denoting the perceived sanction in case of detection, we capture the extrinsic cost of cheating by the expected sanction $S[p(m_{t'}, m_t), s(m_{t'}, m_t)]$. Comparing only the monetary payoff and the extrinsic cost of cheating, the agent will cheat whenever $m_{t'} - m_t > S[p(m_{t'}, m_t), s(m_{t'}, m_t)]$. This inequality illustrates the fundamental trade-off from Becker's (1968) model on the economics of crime: An agent cheats if the benefits of dishonesty outweigh the expected costs.

As discussed in the paper's introduction, the agent's decision may additionally depend on her intrinsic disutility of cheating. For example, a person might have a bad conscience if she realizes that she did not comply with her moral standards. We capture the disutility from not reporting truthfully by adding an intrinsic (psychological) cost of cheating $0 \leq C_i \leq \infty$ to the agent's decision problem. Following Kajackaite and Gneezy (2017), we make the simplifying assumption that C_i is a fixed cost (i.e., it does not depend on the extent of cheating denoted by $t' - t$ and $m_{t'} - m_t$).

Finally, we extend the framework such that it incorporates psychological reactance. Assume the agent faces a situation in which an external request to report truthfully is activated, indicated by $r = 1$; if such a request is not made, then $r = 0$. In the case of an external request, a reactant agent obtains an additional fixed intrinsic utility of cheating $0 \leq R_i \leq \infty$. As

[25] https://doi.org/10.1257/rct.9683.

discussed in the introduction, reactance makes cheating more attractive and reflects the psychological benefit of regaining one's freedom of choice by not reporting truthfully under a request to tell the truth. Note that we allow for heterogeneity in C_i and R_i. Putting the extrinsic and intrinsic costs and benefits of cheating together, the agent will not report truthfully if

$$m_{t'} - m_t - S[p(m_{t'}, m_t), s(m_{t'}, m_t)] - C_i + R_i \cdot 1\{r = 1\} > 0, \qquad \text{(B.1)}$$

where $1\{\cdot\}$ is an indicator function.

Equation (B.1) mirrors the channels through which commitment requests can affect cheating. On the one hand, commitment requests may increase the intrinsic disutility of cheating C_i. On the other hand, reactant agents derive additional intrinsic utility from cheating R_i, if they are requested to commit to truthful reporting. Different forms of commitment requests can thus lead to more or less cheating, depending on how sharply C_i and R_i are shifted.[26]

Table B.1 provides results from linear probability models in which we interact the treatment dummies with indicators for whether a participant belongs to the medium or high tertile of psychological reactance in our sample. Pooling the data from the online and offline setting (Column 1), coefficients for potential interactions are small and statistically insignificant interaction. Analyzing potential heterogeneity separately for the lab (Column 2) and online data (Column 3), we also find no indication for systematic heterogeneity in reactions to our treatments.

[26]Conditional on the setting and the specific form of the declaration of compliance, commitment requests might also change the expected sanction $S[\cdot]$. The the discussion of this topic in the description of the experimental design.

Table B.1: Interactions Between Treatments and Reactance Tertiles

	All Observations (1)	Laboratory (2)	Online (3)
Ethics Treatment	0.024	-0.052	0.076
	(0.073)	(0.095)	(0.134)
Sanction Treatment	-0.036	0.018	-0.060
	(0.073)	(0.101)	(0.120)
Medium Reactance	-0.056	-0.034	-0.093
	(0.077)	(0.095)	(0.113)
High Reactance	-0.010	0.017	-0.007
	(0.078)	(0.102)	(0.132)
Ethics × Medium	-0.036	0.119	-0.014
	(0.106)	(0.132)	(0.173)
Ethics × High	-0.064	-0.048	-0.192
	(0.107)	(0.141)	(0.188)
Sanction × Medium	0.091	-0.044	0.137
	(0.112)	(0.135)	(0.168)
Sanction × High	0.027	-0.033	0.053
	(0.105)	(0.144)	(0.193)
Online	0.017		
	(0.045)		
Lab Konstanz	-0.073	-0.080	
	(0.049)	(0.049)	
N. of obs.	670	406	264
Mean control group	0.322	0.308	0.344

Notes: This table shows linear probability models using as dependent variable an indicator for participants who cheated. Column (1) uses all observations on participants who had a profitable option to cheat. The omitted category for subject pools is Lab Munich. Column (2) uses only observations collected in laboratory sessions (Konstanz and Munich). Column (3) uses only observations collected online (subject pool of laboratory Nuremberg). *Medium Reactance* is an indicator for participants in the second tertile regarding reactance. *High Reactance* is an indicator for participants in the third tertile regarding reactance. The number of observations is slightly lower than in Figure 1 and Table A.1 because some participants did not show up for the online survey after the experimental sessions. We could therefore not elicit these participants' reactance. Robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Online Appendix C. Measuring Psychological Reactance

Psychological Reactance Scale (Hong, 1992)

The following statements concern your general attitudes. Read each statement and please indicate how much you agree or disagree with each statement. If you strongly agree, mark a 5. If you strongly disagree, mark a 1. If the statement is more or less true of you, find the number between 5 and 1 that best describes you. There are no right or wrong answers. Just answer as accurately as possible.

Behavioral and Cognitive Component (De las Cuevas et al., 2014)

1. Regulations trigger a sense of resistance in me.
2. I find contradicting others stimulating.
3. When something is prohibited, I usually think, "That's exactly what I am going to do."
4. I consider advice from others to be an intrusion.
5. Advice and recommendations usually induce me to do just the opposite.
6. I am content only when I am acting of my own free will.
7. I resist the attempts of others to influence me.
8. When someone forces me to do something, I feel like doing the opposite.

Affective Component (De las Cuevas et al., 2014)

9. The thought of being dependent on others aggravates me.
10. I become frustrated when I am unable to make free and independent decisions.
11. It irritates me when someone points out things, which are obvious to me.
12. I become angry when my freedom of choice is restricted.
13. It makes me angry when another person is held up as a role model for me to follow.
14. It disappoints me to see others submitting to standards and rules.