



# Enhanced anonymity in tax experiments does not affect compliance

Christoph Kogler<sup>a,\*</sup>, Jerome Olsen<sup>b,c</sup>, Rebecca I. Bogaers<sup>d</sup>

<sup>a</sup> Tilburg University, School of Social and Behavioral Sciences, Department of Social Psychology, Tilburg, the Netherlands

<sup>b</sup> Max Planck Institute for Research on Collective Goods, Bonn, Germany

<sup>c</sup> University of Vienna, Department Occupational, Economic and Social Psychology, Vienna, Austria

<sup>d</sup> Tilburg University, School of Social and Behavioral Sciences, Scientific Center for Care and Wellbeing, Tilburg, the Netherlands



## ARTICLE INFO

### Article history:

Received 15 October 2019

Revised 10 June 2020

Accepted 23 June 2020

Available online 10 July 2020

### Keywords:

Experiment

Anonymity

Tax compliance

Tax game

Lab procedures

## ABSTRACT

In the domain of classical economic games, it has previously been suggested that deviations from purely rational behavior could be explained by a lack of experimenter-subject anonymity. In fact, some experiments show that contributions and prosocial behavior increase when participants feel observed. In the present study, we investigate whether measures of enhanced anonymity, beyond a conventional standard, are necessary in the particular case of tax behavior experiments. This issue might be pivotal for both the validity and generalizability of existing published studies as well as for designing future studies. We suspect social desirability to be even more relevant in experiments on tax compliance, which often apply a context-rich setting, entailing a strong ethical component. Interestingly, certain common experimental practices reflect potential breaches of anonymity during sign-up, the actual task, and the payment phase. Accordingly, we tested whether (1) tax compliance is higher under conditions of regular anonymity compared to enhanced anonymity, and (2) whether this anonymity manipulation moderates established effects of tax-related parameters, such as audit probability and fine rate. Despite an enhanced perception of anonymity due to our manipulation, we did not observe a difference in relative tax compliance between the regular and enhanced anonymity conditions. Additionally, enhanced anonymity did not interact with the effects of tax rate, audit probability, and fine level on tax compliance. We conclude that commonly used procedures in tax experiments are sufficient to guarantee a satisfactory level of anonymity.

© 2020 The Author(s). Published by Elsevier B.V.  
This is an open access article under the CC BY license.  
(<http://creativecommons.org/licenses/by/4.0/>)

## 1. Introduction

Lab experiments are a key source of knowledge in the social sciences in general (Falk and Heckman, 2009) and are widespread in research on tax behavior (Muehlbacher and Kirchler, 2016). The key strengths of tax experiments are their high internal validity and relative cost-efficiency in studying compliance behavior. Most importantly, in contrast to other widespread empirical research methods (e.g., cross-sectional surveys and administrative data sources) the systematic ma-

\* Corresponding author.

E-mail address: [c.kogler@uvt.nl](mailto:c.kogler@uvt.nl) (C. Kogler).

nipulation (i.e., variation) of relevant factors in lab experiments allows for causal interpretations. At the same time, tax experiments have also been criticized for their assumed low external validity as decision environments in the lab are more artificial and typically lack meaningful real-world consequences.

Considering, for instance, tax-related decisions in private households, it is striking that these decisions are made in private and that they include a strong ethical component. Being or feeling observed can substantially influence performance, decision making, and voicing opinions (see for instance Becker and Marique, 2014; Horwitz and McCaffrey, 2008; Joinson, 1999), so there might be systematic differences between tax decisions in the private household compared to those taken in the lab. Accordingly, one study showed that giving participants a week to make their tax reporting decisions at home significantly decreased tax compliance rates (Cadsby et al., 2006). Potential explanations for this finding include spending more time to make the compliance decision or having contact with other people and maybe even discussing the decision at hand. However, it is also possible that the physical absence of the experimenter accounts for the effect on compliance.

In line with this argument, a number of studies reveal that people behave differently when feeling observed in comparison to feeling unobserved (e.g., Forgas et al., 1980; Froming et al., 1982; Triplett, 1898; van Bommel, van Prooijen, Elffers, and van Lange, 2012; Zajonc, 1965). Often these changes in behavior are driven by strategic motives (e.g., Eagly et al., 1978) and image concerns (e.g., Levitt and List, 2007; Silverman et al., 2014). However, different behavior when feeling observed can also represent spontaneous reactions that do not involve conscious attempts to please the observers (e.g., Noah et al., 2018). Accordingly, people presumably adopt an external perspective of themselves (Hass, 1984) and show a tendency to neglect internal information (Scheier and Carver, 1980; Wicklund and Duval, 1971). This leads to judgments and decisions being biased towards information that is also visible to the potential observers (Steinmetz, Xu, Fishbach, & Zhang, 2016).

These findings are relevant for lab experiments, where participants often engage with the experimenter, for instance when being assigned a computer/cubicle, when asking questions about the procedure, or when receiving their payoff in case of monetary incentives, which is a widespread approach in tax experiments. It would be plausible that such interactions influence compliance decisions. First, they can activate a normative decision context, influencing participants' compliance decisions because they feel more observed and accountable in the presence of the experimenter. Additionally, when monetary incentives are decision-contingent, the experimenter can usually deduce from the final payment amount whether a participant evaded tax or not. Participants who realize this could be hesitant to reveal their true preference of evasion, fearing implicit moral judgement from the experimenter or other participants when receiving their payoff. Second, while participants are usually guaranteed anonymity in the sense that true identities are unknown to the experimenter, they might mistrust the procedures and hold the belief that, for instance, student IDs used for study sign-up or signatures on consent forms and payoff sheets could be linked to collected data and used to reveal their identity. If participants were to hold such concerns, decision outcomes could be biased towards more socially desirable decisions. That is, an enhanced level of compliance and less susceptibility to the experimental manipulations, ultimately resulting in elevated compliance or even ceiling effects.

Being aware of these potentially detrimental effects, some researchers in the field of tax behavior address such concerns in their procedures. For instance, they carefully consider possible breaches of experimenter-subject anonymity (hereafter simply referred to as *anonymity*) and try to ensure an environment of pronounced anonymity in the lab. In this vein, some experimenters explicitly mention that they omit the collection of any real names and assign IDs to participants (e.g., Alm et al., 2010; Blaufus et al., 2016; Choo et al., 2016; Christian and Alm, 2014; Kastlunger et al., 2009; Mittone et al., 2017; Papoutsis et al., 2015; Watrin and Ullmann, 2004; Zhang et al., 2016). Other experimenters refrain from having participants sign a consent form to guarantee anonymity (e.g., Alm et al., 2017; Alm et al., 2010). Furthermore, there are even experimenters applying double-blind payment procedures, for instance, by placing sealed envelopes for each participation ID in the department's administrative office to be collected after the experiment (e.g., Alm et al., 2010; Choo et al., 2016; Christian and Alm, 2014; Mittone et al., 2017; Zhang et al., 2016).

However, although some experiments address this problem with regard to sign-up and payment, they typically do not tackle the potential problem of experimenter-participant interaction. Moreover, it seems that the majority of tax experiments in the lab does not address any of these issues. They either do not explicitly refer to anonymity concerns or only provide a generic statement that anonymity was guaranteed without specifying the respective procedures. Therefore, it often remains unclear whether applied experimental practices are potentially biased due to an insufficient level of anonymity.

Hence, the crucial question is whether measures of enhanced anonymity beyond a common standard are necessary in lab experiments on tax behavior to reduce socially desirable responding. First, this is of practical relevance for experimenters who design new experiments. Second, a clear effect of anonymity on compliance rates might contest existing results of already published experiments that neglected these anonymity concerns.

In the domain of classical game theoretical paradigms, the influence of varying degrees of anonymity has already attracted attention. It has been suggested that deviations from what is considered purely rational behavior, for instance in dictator games, might be explained by a potential lack of complete anonymity. There are in fact studies showing that people are more generous towards another person when they are observed by a third person (e.g., Bull and Gibson-Robinson, 1981; Kurzban, 2001). This is commonly attributed to the finding that cooperative behavior is often sustained by sanctions for those who refuse to act prosocial (Barclay, 2004; Fehr and Gächter, 2002; Milinski et al., 2002; Sylwester and Roberts, 2010; Wedekind and Milinski, 2000). Interestingly, Haley and Fessler (2005) showed that just presenting participants with subtle eye-like stimuli during a dictator game is sufficient to increase contributions. Since participants in this study were explicitly

guaranteed full anonymity, their finding was interpreted as an activation of automatic cognitive mechanisms of detecting social gaze. This could be enough to feel observed and ultimately regulate social behavior in the direction of trying to make a positive impression (Izuma, 2012). In line with this, a number of related studies indicate that eye images can enhance generosity and prosocial behavior in naturalistic settings (Bateson et al., 2006; Ekström, 2012; Ernest-Jones et al., 2011; Francey and Bergmüller, 2012; Powell et al., 2012).

Nevertheless, studies on the potential effects of anonymity on contributions in economic games are somewhat inconclusive. While some studies suggest an influence of different degrees of anonymity on contribution behavior (Franzen and Pointner, 2012; Hoffman et al., 1994; Thielmann et al., 2016), others suggest that this is not the case (Barmettler et al., 2012; Bolton et al., 1998; Bolton and Zwick, 1995).

Irrespective of these ambiguous results, these findings cannot be directly applied to the field of tax research, since there are crucial differences between classical economic games and tax compliance experiments. Tax experiments are often embedded in a context-rich setting. This means that tax language is used (e.g., income, tax rate, audit probability, fine) and the decision entails a stronger ethical component, since participants are assumed to view honesty (= full compliance) as the socially expected behavior (Bruttel and Friehe, 2014; for similar findings in a context not directly related to taxation see Karakostas and Zizzo, 2015). Studies directly comparing tax-specific and context-free language mostly find higher compliance when tax-specific language is used (e.g., Baldry, 1986; Choo, Fonseca, & Myles, 2015; Mittone, 2006; Trivedi and Chung, 2006; Wartick et al., 1999; Webley and Halstead, 1986) with a few opposite findings (e.g., Alm et al., 1992). These are not only absolute differences in reported income, but also interaction effects with tax system variables (Durham et al., 2014). Thus, we believe that it is important to test effects related to anonymity concerns in a setting that applies a tax-specific context. In such a situation, feeling observed or even judged, socially desirable responses should be more likely, especially so when participants do not feel completely anonymous throughout the experiment.

In order to investigate a potential effect of enhanced anonymity on compliance decisions, we conducted a conventional income tax reporting game with varying tax-related parameters over 12 rounds (i.e., tax rate, audit probability, fine rate) where income had to be earned with an effort task and monetary incentives were decision-contingent. As a between-subject factor, we manipulated the degree of anonymity (regular anonymity vs. enhanced anonymity; for details see method section). Our first research question was whether compliance rates vary between different degrees of anonymity. We expected tax compliance to be higher under conditions of regular anonymity compared to enhanced anonymity, based on the assumption that socially desirable responses might become less prevalent with increasing anonymity. Second, we investigated whether our anonymity manipulation moderates the effects of tax-related parameters on tax compliance in the experiment. The underlying rationale is that if differing degrees of anonymity affect the overall level of compliance, it could be the case that deterrence effects (i.e., effects of audit probabilities and fine levels) might work differently. For instance, stronger effects of social desirability in the regular anonymity condition could mitigate the effect of deterrence on compliance compared to a condition of enhanced anonymity.

## 2. Method

### 2.1. Participants

The sample comprised 131 students. Mean age was  $M = 21.55$  ( $SD = 2.56$ ) and 79 (60.3%) were female. All participants were enrolled in the Psychology Bachelor Program at Tilburg University and were invited to participate in the study to earn course credits. They were informed about additional performance-based monetary payment. About half of the participants were international students (54.6%) and the other half Dutch (45.4%).

The sample size resulted from the available resources and recruitment success during the two test weeks. The obtained sample size provided a power of 0.80 to detect an effect as small as Cohen's  $d = 0.49$  in terms of a mean difference between the two anonymity conditions (two-sided  $t$ -test with  $\alpha = 0.05$ ).<sup>1</sup>

### 2.2. Materials

#### 2.2.1. Experiment

The study consisted of a repeated rounds tax experiment including 12 compliance decisions. Within a mixed-design, two different degrees of anonymity were manipulated between participants (regular anonymity versus enhanced anonymity; see Procedure section) and three within-subject factors were fully permuted over the 12 rounds in a randomly determined fixed order (tax-rate, 2 levels [20% or 40%]; audit probability, 3 levels [5%, 15%, or 25%]; fine rate, 2 levels [evaded tax plus a fine of 0.5 times the evaded amount or evaded tax plus a fine of 1.5 times the evaded amount]). Hence, the scheme of the parameters tax rate, audit probability, and fine rate over the rounds was identical for all participants in both conditions.

All instructions and descriptions used context-rich language (i.e., a strong tax frame), in order to emphasize that individuals made compliance decisions within a normative tax frame where compliance can be seen as the expected behavior. For

<sup>1</sup> Note that this might be a conservative estimate as the actual design consists of 12 repeated measures and was modeled using a linear mixed effects model.

instance, participants were informed that “the tax rate is the percentage of your income you have to pay as tax” or that the fine is “what you have to pay in case you did not pay all of your tax due and get audited”.

In each round, individuals received a guaranteed income of 1000 experimental currency units (ECU) and had the opportunity to earn up to 1000 ECU in addition, based on performance in a real-effort slider task. Accordingly, income could vary between 1000 and 2000 ECU ( $M = 1648$ ,  $SD = 203$ ). After earning their income in each round, individuals received information about the tax rate, audit probability, and fine rate, and had to decide how much of the tax due to pay. The audit mechanism was implemented in accordance with the announced deterrence parameters and feedback on whether they were audited (and on potential fines) was provided directly after each round of tax filing.

Tax decisions were incentivized (exchange rate: 1 Euro = 300 ECU). Remuneration was based on one randomly selected round plus an additional show-up fee of 1 Euro. Average remuneration was  $M = 5.38$  Euro ( $SD = 1.26$ ).

Before the actual experiment, participants were presented with three exemplary tax rounds where they had to indicate the resulting net income in ECU to check whether they understood the instructions. Furthermore, participants also faced two trial rounds to make themselves familiar with the effort task and the software in general. An audit took place in the second of these two rounds.

### 2.2.2. Post-experimental questionnaire

After the 12 rounds of the experiment, participants had to fill in a post-experimental questionnaire. The questionnaire consisted of six blocks. In all blocks, items had to be answered on a Likert-type scale ranging from 1 = *lowest agreement* to 5 = *highest agreement*.

First, they were presented with eight statements measuring tax attitudes in general (i.e., the moral obligation to pay taxes and to support the principles of taxation). These items were taken from the Motivational Postures *Commitment* subscale (Braithwaite, 2003; e.g., “Paying tax is the right thing to do.”).

Second, individuals had to provide answers to two items on personal norms (e.g., “Do you think it is acceptable to evade some tax?”) and to two questions about social norms toward paying tax (e.g., “Do most people think it is acceptable to evade some tax?”). These items were constructed for the purpose of this study.

Third, we collected basic socio-demographic information. On the same page, we asked whether participants read all instruction carefully and asked to indicate whether they understood the instructions.

Fourth, individuals had to provide answers to three multiple-choice attention checks for the used tax-related parameters (i.e., “What were the different fine levels in the tax game?”) with one of the four provided options being correct.

Fifth, participants were presented with three items measuring perceived anonymity (1) during the sign-up phase for the study, (2) during the actual experiment in the lab, and (3) in the payment stage (e.g., “The payment procedure guarantees my anonymity”; for details of the manipulation see next section).

Finally, we asked participants about their perceptions and motives when deciding on whether to comply or evade. This allows for exploring whether decisions were perceived as paying tax or rather as a game (e.g., “I perceived the decisions as playing a game”), and whether self-stated motives were payoff maximization or rather tax honesty (e.g., “My main goal was to maximize my payoff”). All materials were presented in English (see the OSF repository for screenshots of the instructions and items; <https://osf.io/6gu2z>).

### 2.3. Procedure

The manipulation of regular anonymity versus enhanced anonymity was implemented in three different procedural steps of the study: (1) during the sign-up for the study, (2) during the actual experimental task, and (3) in the payment phase.

In the regular anonymity condition,<sup>2</sup> individuals signed up for the study indicating their student ID number and had to choose a time slot. When arriving at the lab to participate, individuals were assigned to a cubicle by the experimenter in person. The lab consisted of multiple cubicles with closable doors. These doors had to stay open throughout the experimental task and the experimenter answered potential questions in person<sup>3</sup> (“Please leave the doors of your cubicle open during the study. In case it is closed, please open it now.”). After completion of the post-experimental questionnaire, the payoff was displayed on the computer screen. The experimenter went to each cubicle and paid each participant in person. Participants were informed about the payment procedure at the beginning of the experiment (“The experimenter will hand you the money at the end of the study.”). As detailed, in the regular anonymity condition participants interacted with the experimenter several times in person and they were aware that the experimenter will see their final payoff.

In the enhanced anonymity condition, individuals also signed up using their student ID number. Subsequently, they received an email stating that their student ID number would not be used in the study and that there would be no personal interaction with an experimenter, both to ensure complete anonymity. In the same email, participants were instructed to select one of the available cubicles and then close the cubicle door once they arrive at the lab. A large poster in the arrival room of the lab also displayed this information in case someone did not remember these instructions and they were also mentioned in the experimental instructions (e.g., “Please make sure the door of your cubicle is closed during the study.

<sup>2</sup> The normative term *regular* reflects that this is close to how we would typically conduct a tax experiment in the lab.

<sup>3</sup> We left the doors open in this experimental condition to account for the fact that in many labs cubicles are not closable.



**Fig. 1.** Mean experienced anonymity during the sign-up phase, the actual experiment, and when receiving payment by condition. Error bars represent 95% confidence intervals.

In case it is currently open, please close it now.”). Additionally, individuals were informed that there was an experimenter present in the control room who could be contacted via an intercom system in case of any problems, but they would not see this person otherwise, as there were blinds on all windows of the control room (“...There will be an experimenter present in the control room in case you run into any problems, but if there are no problems you will not have any contact with the experimenter to ensure anonymity...”). To ensure full anonymity during payment, we placed a sticker with a random number-letter code in each cubicle. Individuals were instructed to enter this code at the beginning of the experimental task and to take the code with them afterwards. From 30 min after the completion of the study onward, they could collect an envelope marked with their random number-letter code containing their remuneration in the arrival room of the lab.

Comparing the two conditions, during the sign-up phase participants in the enhanced anonymity condition entered their student ID number, but were informed that it would not be used for any other purpose than crediting mandatory lab hours, whereas participants in the regular anonymity condition did not receive such specific information. During the experiment, individuals in the enhanced anonymity condition did not interact with the experimenter at all and sat in cubicles they selected with closed doors. In the regular anonymity condition, individuals were assigned by the experimenter and the door remained open. During payment, participants in the enhanced anonymity received their remuneration based on unmatchable random codes and without any interaction with the experimenter. Participants in the regular anonymity condition received their money from the experimenter in person. Note that participants in both conditions were fully informed about these procedures before making their tax compliance decisions in the experiment and completing the post-experimental questionnaire.

Sessions of one condition were run for a whole day. To rule out potential date effects on condition assignment, we ran the experiment over the course of two weeks, counterbalancing the weekdays per condition.

Data and R code can be downloaded from <https://osf.io/6gu2z>.

### 3. Results

The results are divided into three sections. First, we tested whether the manipulation of regular versus enhanced anonymity was successful. Second, we ran three regression models to test the effects of the anonymity manipulation, the tax-related parameters, and their interaction on tax compliance. Third, we conducted further exploratory analyses.

#### 3.1. Manipulation check

To test whether the manipulation of regular versus enhanced anonymity was successful, we ran a MANOVA with the anonymity condition as independent variable and the three manipulation check item scores as dependent variables. Fig. 1 shows the perceived anonymity by condition for each of the three items. Results indicate that overall experienced anonymity was higher in the enhanced than in the regular anonymity condition,  $F(3, 127) = 7.34, p < .001, \eta^2 = 0.15$ , indicating a successful manipulation of anonymity.

Univariate analyses by item revealed that the two conditions did not differ significantly from each other in the extent of perceived anonymity during the sign-up phase,  $F(1, 129) = 2.38, p = .126, \eta^2 = 0.02$ , but during the experiment,  $F(1, 129) = 8.31, p = .005, \eta^2 = 0.06$ , and when receiving the payoff,  $F(1, 129) = 21.99, p < .001, \eta^2 = 0.15$ . Altogether, participants in the enhanced anonymity condition felt more anonymous.

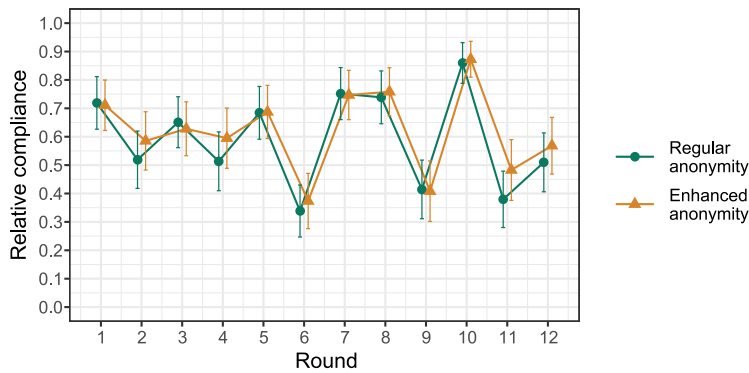
#### 3.2. Tax compliance

We ran three linear mixed effects models with relative tax compliance as dependent variable. We entered a random intercept for individuals in all three models to account for the repeated measures data structure ( $N = 131$ , leading to 1572 observations in 12 rounds). Model 1 included only the anonymity manipulation as a fixed effect, and Model 2 included only the tax-related parameters as fixed effects. In Model 3, we entered the anonymity manipulation, the tax-related parameters, and the interactions of these two as fixed effects. All models are presented in Table 1.

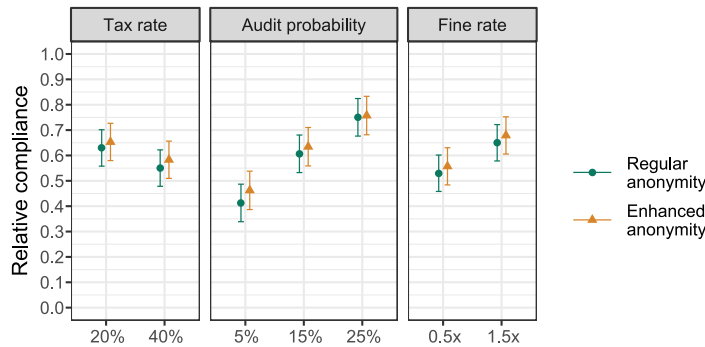
**Table 1**  
Linear mixed effects models with relative tax compliance as dependent variable.

Variables	Relative tax compliance								
	Model 1			Model 2			Model 3		
	B	SE	p	B	SE	p	B	SE	p
Intercept	0.59	0.03	< .001	0.41	0.03	< .001	0.39	0.04	< .001
Anonymity	0.03	0.05	.569				0.04	0.06	.439
Tax rate				−0.07	0.01	< .001	−0.08	0.02	< .001
Audit probability 15%				0.18	0.02	< .001	0.19	0.02	< .001
Audit probability 25%				0.32	0.02	< .001	0.34	0.02	< .001
Fine level				0.12	0.01	< .001	0.12	0.02	< .001
Anonymity x Tax rate							0.01	0.03	.752
Anonymity x Audit prob. 15%							−0.02	0.04	.538
Anonymity x Audit prob. 25%							−0.04	0.04	.225
Anonymity x Fine level							0.00	0.03	.953

Note.  $N_{\text{individuals}} = 131$  over 12 rounds;  $N_{\text{observationsTotal}} = 1572$ . Anonymity was coded with 0 = regular anonymity and 1 = enhanced anonymity. Tax rate was coded with 0 = 20% and 1 = 40%. Audit probability was dummy coded with 5% as reference category. Fine level was coded with 0 = payback plus 0.5x evaded amount and 1 = payback plus 1.5x evaded amount.



**Fig. 2.** Mean relative tax compliance over the 12 rounds by condition. Error bars represent 95% confidence intervals.



**Fig. 3.** Mean relative tax compliance for all levels of the tax-related parameters by condition. Error bars represent 95% confidence intervals. These were extracted from linear mixed effects models with a random intercept for individuals to account for the dependence in the data.

**3.2.1. Model 1: effect of anonymity on tax compliance**

The manipulation of regular versus enhanced anonymity had no significant effect on tax compliance (see Table 1). Fig. 2 reveals that compliance levels were close to congruent. Hence, we found no indication that tax compliance decisions in the lab change under a condition of enhanced anonymity.

To get an estimate of the standardized mean difference between the two conditions, we compared mean relative compliance over the twelve rounds between the two anonymity conditions with a *t*-test,  $t(128.77) = -0.56, p = .578$ . The standardized effect size estimate was Cohen's  $d = -0.10$ , 95% CI  $[-0.44, 0.25]$ .

**3.2.2. Model 2: effects of tax-related parameters on tax compliance**

Compliance was influenced by all three tax-related parameters (see Table 1 and Fig. 3). Individuals' relative compliance was significantly lower tax in rounds with a tax rate of 40% compared to rounds of 20%. The effects of audit probability and



fine rate indicated a clear deterrence effect. Compliance was significantly higher when audits were more likely. Comparing rounds with 5% to rounds with 15% audit probability, relative compliance increased by 18 percentage points. Moving from 15% to 25% increased relative compliance by another 14 percentage points. A fine rate of 1.5 times the evaded amount increased relative compliance by 12 percentage points compared to a fine rate of 0.5.

### 3.2.3. Model 3: Effects of anonymity, tax-related parameters, and their interactions on tax compliance

Entering the interaction terms between the anonymity manipulation and each tax-related parameter did not affect the previously described effects (see Table 1). None of the interaction terms was significant. Hence, the anonymity manipulation did not moderate the effectiveness of any of the tax-related parameters (see Fig. 3).

## 3.3. Exploratory results

### 3.3.1. Instruction and attention checks

The data contained multiple indicators of individuals' understanding of and attention to the instructions: participants were asked to (1) report the correct net income of three presented tax decisions after reading the instructions, (2) correctly recall the presented tax rates, audit probabilities, and fine rates after the experimental task, (3) state whether they read the instructions carefully, and (4) indicate whether they understood the instructions. We tested whether there is an effect of the anonymity manipulation on compliance among participants who stated to understand the instructions or passed the check items in separate models, respectively. Again, there was no effect of anonymity when controlling for any of these data quality indicators ( $p$ -values of anonymity on condition for the four models:  $p = .459$ ,  $p = .623$ ,  $p = .511$ ,  $p = .633$ ).

### 3.3.2. Did the anonymity manipulation influence individuals' perceptions and motives?

After the experiment, participants were asked to recall their perceptions and motives during the task. Specifically, they had to indicate whether they perceived the decisions in the experiment rather as paying taxes or as playing a game, whether they thought of any moral aspects regarding their decisions, and whether they wanted to maximize their payoff or rather pay their taxes due honestly. We assumed that perceptions of playing a game, lower levels of moral considerations, and payoff maximization motives might be more prevalent in the enhanced anonymity condition. These assumptions were not supported by the data as suggested by a MANOVA with the anonymity condition as independent variable and the five motive items as dependent variables,  $F(5, 125) = 0.59$ ,  $p = .706$ . Univariate results further supported this exploratory finding.

## 4. Discussion

Our study tested whether tax compliance in a classical experimental decision setting varies between different degrees of anonymity. Specifically, we investigated whether compliance is more pronounced under conditions of regular anonymity compared to enhanced anonymity, and whether these different degrees of anonymity moderate the effects of prominent determinants of tax compliance. Our results do not support these assumptions.

The absence of a significant interaction between the anonymity manipulation and any of the tax-related parameters is reassuring, as it does not challenge published findings from experiments that do not employ such pronounced anonymity measures, as for instance complete absence of experimenter-subject interactions. Our findings – in general and especially with regard to the payment procedure – are in line with a recent study that shows that having participants enter their names during the experiment for receipt preparation does not lead to more socially desirable answers (Wolff, 2019).

One important conclusion of our study is that past experiments investigating deterrence effects are thus not likely to be considerably influenced or biased by social desirability. This also means that often observed high compliance rates in tax experiments cannot be criticized for just being artifacts due to a lack of anonymity. Hence, on a more general level, this could be interpreted as a fundamental difference between experiments in the context of tax behavior and economic games, as for instance the dictator game, where some studies suggest that perceived anonymity influences the level of contributions (e.g., Bull and Gibson-Robinson, 1981; Franzen and Pointner, 2012; Haley and Fessler, 2005; Hoffman et al., 1994; Kurzban, 2001; Thielmann et al., 2016). One potential explanation for this discrepancy is that our setting (as usual in a standard tax game) did not rely on a public goods structure. Employing such a structure, where the payoff of other participants is directly influenced by own contributions, could increase the influence of anonymity on contribution behavior.

Besides its significance for designing and interpreting experimental studies on tax compliance, our study also contributes to the more general issue of whether studies on cheating and ethical decision making could be biased due to minor anonymity breaches. Experimental approaches like the popular die-under-cup paradigm (Shalvi et al., 2011) have been designed in order to study cheating in a setting where it is transparently clear to participants that they cannot be observed by the experimenter. Comparing such an approach to rather traditional lab experiments to study unethical behavior, our results seem to indicate that the differences might be negligible, as long as participants perceive the level of anonymity as sufficient, keeping in mind that in both our conditions participants did not indicate anonymity to be low or insufficient.

One limitation of the present study is that we only tested the moderating role of anonymity for the effects of classical tax-related parameters on compliance. It remains an open question whether anonymity has an influence on, for instance, effects of fairness or social norms, although we have no evidence that social desirability would play a stronger role in such a context.

Besides that, one could argue that our sample size was not large enough to detect smaller difference between the two conditions. The sample size provided power of 0.80 to detect an effect in the range of Cohen's  $d = 0.50$ . Admittedly, a larger sample would have been preferable to obtain a higher chance to detect a potentially smaller effect. However, the obtained effect estimate (Cohen's  $d = -0.10$ , 95% CI  $[-0.44, 0.25]$ ) suggests that any effect is likely to be in the range of a small effect. Hence, we are confident to infer from the data that there are no large effects of regular versus enhanced anonymity.

We must emphasize that our study compared a situation of regular anonymity against one of enhanced anonymity. This means that we did not have a *low* anonymity condition in absolute terms, which would reflect severe formal breaches of anonymity. Absolute anonymity perceptions are also quite high in the enhanced anonymity condition, as, for instance, identifying individuals' identities in the experimental data obtained was not possible in either condition. However, we do observe a successful manipulation (in terms of our operationalization) with substantial differences in self-stated experienced anonymity between the two conditions, especially for experienced anonymity during the experimental task and for the payment procedure ( $\eta^2 = 0.06$  and  $\eta^2 = 0.15$ , respectively), most likely due to the interaction with the experimenter. Hence, our study tested a small, but practically important difference in experimental design.

We do believe that general anonymity concerns are relevant and our results should not be interpreted as a justification to neglect the importance of anonymity in experimental research. However, they also provide evidence that often-used standard procedures in tax experiments are sufficient to guarantee a satisfactory level of anonymity. Thus, we are confident that the present study contributes to the field of tax research by offering useful insights into how to design experiments and especially in helping to evaluate existing findings in the literature.

### Declaration of Competing Interest

None.

### Acknowledgement

The first author received funding from the [European Research Council \(ERC\)](#) under the European Union's Horizon 2020 research and innovation program (grant agreement No. [798824](#)) for the drafting of this paper.

### References

- Alm, J., Bloomquist, K.M., McKee, M., 2017. When you know your neighbour pays taxes: information, peer effects, and tax compliance. *Fisc Stud* 38, 587–613. <https://doi.org/10.1111/1475-5890.12111>.
- Alm, J., Cherry, T., Jones, M., McKee, M., 2010. Taxpayer information assistance services and tax compliance behavior. *J Econ Psychol* 31 (4), 577–586. doi:[10.1016/j.joep.2010.03.018](https://doi.org/10.1016/j.joep.2010.03.018).
- Alm, J., McClelland, G.H., Schulze, W.D., 1992. Why do people pay taxes. *J Public Econ* 48, 21–38.
- Baldry, J.C., 1986. Tax evasion is not a gamble. *Econ Lett* 22, 333–335.
- Barclay, P., 2004. Trustworthiness and competitive altruism can also solve the “tragedy of the commons. *Evol. Hum. Behav.* 25, 209–220. <https://doi.org/10.1016/j.evolhumbehav.2004.04.002>.
- Barmettler, F., Fehr, E., Zehnder, C., 2012. Big experimenter is watching you! Anonymity and prosocial behavior in the laboratory. *Games Econ Behav* 75, 17–34. <https://doi.org/10.1016/j.geb.2011.09.003>.
- Bateson, M., Nettle, D., Roberts, G., 2006. Cues of being watched enhance cooperation in a real-world setting. *Biol. Lett.* 2, 412–414. <https://doi.org/10.1098/rsbl.2006.0509>.
- Becker, T.E., Marique, G., 2014. Observer effects without demand characteristics: an inductive investigation of video monitoring and performance. *J Bus. Psychol.* 29, 541–553. <https://doi.org/10.1007/s10869-013-9338-1>.
- Blaufus, K., Hundsdorfer, J., Jacob, M., Sünwoldt, M., 2016. Does legality matter? The case of tax avoidance and evasion. *J. Econ. Behav. Organ.* 127, 182–206. <https://doi.org/10.1016/j.jebo.2016.04.002>.
- Bolton, G.E., Katok, E., Zwick, R., 1998. Dictator game giving: rules of fairness versus acts of kindness. *Int. J. Theory* 27, 269–299. <https://doi.org/10.1007/s001820050072>.
- Bolton, G.E., Zwick, R., 1995. Anonymity versus punishment in ultimatum bargaining. *Games Econ. Behav.* 10, 95–121.
- Braithwaite, V., 2003. Taxing democracy: understanding tax avoidance and evasion. In: Braithwaite, V. (Ed.), *Dancing With Tax authorities: Motivational postures and Non-Compliant Actions*. Ashgate, Aldershot, UK, pp. 15–39.
- Bruttel, L., Friehe, T., 2014. On the path dependence of tax compliance. *Eur. Econ. Rev.* 65, 90–107. <https://doi.org/10.1016/j.eurocorev.2013.11.002>.
- Bull, R., Gibson-Robinson, E., 1981. The influences of eye-gaze, style of dress and locality on the amounts of money donated to a charity. *Hum. Relat.* 34, 895–905. doi:[10.1177/001872678103401005](https://doi.org/10.1177/001872678103401005).
- Cadsby, C.B., Maynes, E., Trivedi, V.U., 2006. Tax compliance and obedience to authority at home and in the lab: a new experimental approach. *Exp. Econ.* 9, 343–359. <https://doi.org/10.1007/s10683-006-7053-8>.
- Choo, C.Y.L., Fonseca, M.A., Myles, G.D., 2016. Do students behave like real taxpayers in the lab? Evidence from a real effort tax compliance experiment. *J. Econ. Behav. Organ.* 124, 102–114. <https://doi.org/10.1016/j.jebo.2015.09.015>.
- Christian, R.C., Alm, J., 2014. Empathy, sympathy, and tax compliance. *J. Econ. Psychol.* 40, 62–82. doi:[10.1016/j.joep.2012.10.001](https://doi.org/10.1016/j.joep.2012.10.001).
- Durham, Y., Manly, T.S., Ritsema, C., 2014. The effects of income source, context, and income level on tax compliance decisions in a dynamic experiment. *J. Econ. Psychol.* 40, 220–233. <https://doi.org/10.1016/j.joep.2012.09.012>.
- Eagly, A.H., Wood, W., Chaiken, S., 1978. Causal inferences about communicators and their effect on opinion change. *J. Pers. Soc. Psychol.* 36, 424–435. <https://doi.org/10.1037/0022-3514.36.4.424>.
- Ekström, M., 2012. Do watching eyes affect charitable giving? Evidence from a field experiment. *Exp. Econ.* 15, 530–546. <https://doi.org/10.1007/s10683-011-9312-6>.
- Ernest-Jones, M., Nettle, D., Bateson, M., 2011. Effects of eye images on everyday cooperative behavior: a field experiment. *Evol. Hum. Behav.* 32, 172–178. <https://doi.org/10.1016/j.evolhumbehav.2010.10.006>.
- Falk, A., Heckman, J.J., 2009. Lab experiments are a major source of knowledge in the social sciences. *Science* 326, 535–538. <https://doi.org/10.1126/science.1168244>.
- Fehr, E., Gächter, S., 2002. Altruistic punishment in humans. *Nature* 415, 137–140. <https://doi.org/10.1038/415137a>.



- Forgas, J.P., Brennan, G., Howe, S., Kane, J.F., Sweet, S., 1980. Audience effects on squash players' performance. *J. Soc. Psychol.* 111, 41–47. <https://doi.org/10.1080/00224545.1980.9924271>.
- Francey, D., Bergmüller, R., 2012. Images of eyes enhance investments in a real-life public good. *PLoS ONE* 7, e37397. doi:10.1371/journal.pone.0037397. <https://doi.org/>
- Franzen, A., Pointner, S., 2012. Anonymity in the dictator game revisited. *J. Econ. Behav. Organ.* 81, 74–81. <https://doi.org/10.1016/j.jebo.2011.09.005>.
- Froming, W.J., Walker, G.R., Lopyan, K.J., 1982. Public and private self-awareness: when personal attitudes conflict with societal expectations. *J. Exp. Soc. Psychol.* 18, 476–487.
- Haley, K.J., Fessler, D.M.T., 2005. Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evol. Hum. Behav.* 26, 245–256. <https://doi.org/10.1016/j.evolhumbehav.2005.01.002>.
- Hass, R.G., 1984. Perspective taking and self-awareness: drawing an E on your forehead. *J. Pers. Soc. Psychol.* 46, 788–798.
- Hoffman, E., McCabe, K., Shachat, K., Smith, V., 1994. Preferences, property rights, and anonymity in bargaining games. *Games Econ. Behav.* <https://doi.org/10.1006/game.1994.1056>.
- Horwitz, J.E., McCaffrey, R.J., 2008. Effects of a third party observer and anxiety on tests of executive function. *Arch. Clin. Neuropsychol.* 23, 409–417. <https://doi.org/10.1016/j.acn.2008.02.002>.
- Izuma, K., 2012. The social neuroscience of reputation. *Neurosci. Res.* 72, 283–288. <https://doi.org/10.1016/j.neures.2012.01.003>.
- Jonson, A., 1999. Social desirability, anonymity, and internet-based questionnaires. *Behav. Res. Methods, Instrum. Comput.* 31 (3), 433–438.
- Karakostas, A., Zizzo, D.J., 2015. Compliance and the power of authority. *J. Econ. Behav. Organ.* 124, 67–80. <https://doi.org/10.1016/j.jebo.2015.09.016>.
- Kastlunger, B., Kirchler, E., Mittone, L., Pitters, J., 2009. Sequences of audits, tax compliance, and taxpaying strategies. *J. Econ. Psychol.* 30, 405–418. <https://doi.org/10.1016/j.joep.2008.10.004>.
- Kurzban, R., 2001. The social psychophysics of cooperation: nonverbal communication in a public goods game. *J. Nonverbal. Behav.* 25, 241–259. <https://doi.org/10.1023/a:1012563421824>.
- Levitt, S.D., List, J.A., 2007. What do laboratory experiments measuring social preferences reveal about the real world. *J. Econ. Perspect.* 21, 153–174. <https://doi.org/10.1257/jep.21.2.153>.
- Milinski, M., Semmann, D., Krambeck, H.J., 2002. Reputation helps solve the 'tragedy of the commons. *Nature* 415, 424–426. <https://doi.org/10.1038/415424a>.
- Mittone, L., 2006. Dynamic behaviour in tax evasion: an experimental approach. *J. Socio-Econ.* 35, 813–835. <https://doi.org/10.1016/j.socec.2005.11.065>.
- Mittone, L., Panebianco, F., Santoro, A., 2017. The bomb-crater effect of tax audits: beyond the misperception of chance. *J. Econ. Psychol.* 61, 225–243. <https://doi.org/10.1016/j.joep.2017.04.007>.
- Muehlbacher, S., Kirchler, E., 2016. Taxperiments: about the external validity of laboratory experiments in tax compliance research. *Die Betriebswirtschaft* 76, 7–19.
- Noah, T., Schul, Y., Mayo, R., 2018. Thinking of oneself as an object of observation reduces reliance on metacognitive information. *J. Exp. Psychol.: Gen.* 147 (7), 1023–1042. <https://doi.org/10.1037/xge0000440>.
- Papoutsis, G.S., Nayga, R.M., Lazaridis, P., Drichoutis, A.C., 2015. Fat tax, subsidy or both? The role of information and children's pester power in food choice. *J. Econ. Behav. Organ.* 117, 196–208. <https://doi.org/10.1016/j.jebo.2015.06.011>.
- Powell, K.L., Roberts, G., Nettle, D., 2012. Eye images increase charitable donations: evidence from an opportunistic field experiment in a supermarket. *Ethology* 118, 1–6. <https://doi.org/10.1111/eth.12011>.
- Scheier, M.F., Carver, C.S., 1980. Private and public self-attention, resistance to change, and dissonance reduction. *J. Pers. Soc. Psychol.* 39, 390–405. <https://doi.org/10.1037/0022-3514.39.3.390>.
- Shalvi, S., Dana, J., Handgraaf, M.J.J., Dreu, C.K.W.De, 2011. Justified ethicality: observing desired counterfactuals modifies ethical perceptions and behavior. *Organ. Behav. Hum. Decis. Process.* 115, 181–190. <https://doi.org/10.1016/j.obhdp.2011.02.001>.
- Silverman, D., Slemrod, J., Uler, N., 2014. Distinguishing the role of authority "in" and authority "to. *J. Public. Econ.* 113, 32–42. <https://doi.org/10.1016/j.jpubeco.2014.02.003>.
- Sylwester, K., Roberts, G., 2010. Cooperators benefit through reputation-based partner choice in economic games. *Biol. Lett.* 6, 659–662. <https://doi.org/10.1098/rsbl.2010.0209>.
- Thielmann, I., Heck, D.W., Hilbig, B.E., 2016. Anonymity and incentives: an investigation of techniques to reduce socially desirable responding in the Trust Game. *Judgm. Decis. Mak.* 11, 527–536.
- Triplet, N., 1898. The dynamogenic factors in pacemaking and competition. *Am. J. Psychol.* 9 (4), 507–533.
- Trivedi, V.U., Chung, J.O.Y., 2006. The impact of compensation level and context on income reporting behavior in the laboratory. *Behav. Res. Account.* 18, 167–183.
- van Bommel, M., van Prooijen, J.-W., Elffers, H., van Lange, P.A.M., 2012. Be aware to care: public self-awareness leads to a reversal of the bystander effect. *J. Exp. Soc. Psychol.* 48, 926–930. <https://doi.org/10.1016/j.jesp.2012.02.011>.
- Wartick, M.L., Madeo, S.A., Vines, C.C., 1999. Reward dominance in tax-reporting experiments: the role of context. *J. Am. Tax. Assoc.* 21, 20–31.
- Watrin, C., Ullmann, R., 2004. Comparing direct and indirect taxation: the influence of framing on tax compliance. *Eur. J. Comp. Econ.* 5, 33–56. Retrieved from <http://eaces.liuc.it/18242979200801/182429792008050102.pdf>.
- Webley, P., Halstead, S., 1986. Tax evasion on the micro: significant simulations or expedient experiments. *J. Interdisc. Econ.* 1, 87–100.
- Wedekind, C., Milinski, M., 2000. Cooperation through image scoring in human. *Science* 288, 850–853. <https://doi.org/10.1126/science.288.5467.850>.
- Wicklund, R.A., Duval, S., 1971. Opinion change and performance facilitation of objective as a result. *J. Exp. Soc. Psychol.* 7, 319–342. [https://doi.org/10.1016/0022-1031\(71\)90032-1](https://doi.org/10.1016/0022-1031(71)90032-1).
- Wolff, I., 2019. The reliability of questionnaires in laboratory experiments: what can we do. *J. Econ. Psychol.* 74. <https://doi.org/10.1016/j.joep.2019.102197>.
- Zajonc, R.B., 1965. Social facilitation. *Science* 149, 269–274. <https://doi.org/10.1126/science.149.3681.269>.
- Zhang, N., Andrighetto, G., Ottone, S., Ponzano, F., Steinmo, S., 2016. "Willing to pay?" Tax compliance in Britain and Italy: an experimental analysis. *PLoS ONE* 11, e0150277. <https://doi.org/10.1371/journal.pone.0150277>.