



# Delayed feedback on tax audits affects compliance and fairness perceptions



Christoph Kogler<sup>a,\*</sup>, Luigi Mittone<sup>b</sup>, Erich Kirchler<sup>a</sup>

<sup>a</sup> University of Vienna, Faculty of Psychology, Institute of Applied Psychology: Work, Education and Economy, Universitaetsstrasse 7, 1010 Vienna, Austria

<sup>b</sup> University of Trento, CEEL – Computable and Experimental Economics Laboratory, Via Calepina, 14, 38122 Trento, Italy

## ARTICLE INFO

### Article history:

Received 16 June 2014  
Received in revised form 15 April 2015  
Accepted 28 October 2015  
Available online 17 November 2015

### JEL classification:

H26

### Keywords:

Tax compliance  
Delayed feedback  
Procedural justice  
Audit probability  
Tax morale

## ABSTRACT

In the present study we explore the conflicting finding that delayed feedback on tax audits apparently results in higher tax compliance, although delaying feedback is associated with lower perceptions of procedural fairness. In a repeated rounds design the timing of feedback (delayed vs. immediate) is investigated in combination with a reduction of wealth in some periods, presented either as a rather unfair intervention of the authorities, or due to a comparatively neutral manipulation. The results reveal a strong impact of timing of feedback on tax compliance: participants in conditions of delayed feedback show significantly higher compliance than those in conditions of immediate feedback. In addition, participants receiving delayed feedback are more likely to state that the probability of audit is high as well as that fines in case of detection are severe, but on the other hand perceive the timing of feedback and the authorities as more unfair. No main effect of the reason for the wealth reduction could be identified. The finding that the difference in compliance between delayed and immediate feedback on tax audits develops over time supports the assumption that in decisions based on experience the probabilities of rare events are underweighted, while in decisions based on descriptions people make choices as if they overweight probabilities of rare events.

© 2015 Elsevier B.V. All rights reserved.

## 1. Introduction

Experimental research on tax behavior typically suffers from one important discrepancy in comparison to tax paying in reality: In experiments, audits usually occur directly after filing taxes, and feedback on the consequences of an audit is given immediately. In reality, audits happen within a much longer period of time. In Austria, for instance, taxpayers can be audited up to ten years after filing taxes. This difference of time lags between filing and audit may play a crucial role with regard to the external validity of experimental results, as well as for the determination of feedback intervals on tax audits in reality. Indeed, there is evidence, that a longer interval between taxpaying and information, whether an audit takes place, enhances compliance (Muehlbacher et al., 2012). Thus, this time-lag effect might be explained by negative feelings associated with unfinished businesses (cf., Zeigarnik, 1927), a

\* Corresponding author at: Department of Applied Psychology: Work, Education and Economy, Faculty of Psychology, University of Vienna, Universitaetsstrasse 7, A-1010 Vienna, Austria. Tel.: +43 1 4277 47335; fax: +43 1 4277 47339.

E-mail addresses: [christoph.kogler@univie.ac.at](mailto:christoph.kogler@univie.ac.at) (C. Kogler), [luigi.mittone@economia.unitn.it](mailto:luigi.mittone@economia.unitn.it) (L. Mittone), [erich.kirchler@univie.ac.at](mailto:erich.kirchler@univie.ac.at) (E. Kirchler).

preference to bring-forward unpleasant events to minimize the dread experienced while waiting (Badia et al., 1966; Loewenstein, 1987) and a prolonged period of uncertainty. Thus, anticipating the dread related to cheating can be avoided when filing an honest tax return in order to prevent the uncertainty associated with a possible audit after evading taxes.

Another explanation for a potential effect of delayed information on tax audits stems from the field of research on judgment and decision making. Accordingly, decisions from description and decisions based on experience can lead to clearly different choice behavior. While in decisions based on description people make choices as if they overweight probabilities of rare events (cf. Kahneman and Tversky, 1979), in decisions based on experience the probabilities of rare events are said to be underweighted (Hertwig et al., 2004; Hertwig and Erev, 2009). With regard to the tax context this would imply that decisions with delayed feedback on audits are predominantly based on descriptions of the experimental situation and especially the occurrence of audits, whereas in decisions with immediate feedback a major influence of experience should prevail. Thus, delayed feedback should result in higher tax compliance due to overweighting of the probability of a tax audit, and, respectively, immediate feedback lead to lower compliance due to underweighting of audit probability.

The assumption that delayed feedback on audits leads to a higher level of tax compliance is contrasting the fact that in the context of inter-organizational relations, the provision of timely feedback increases perceived procedural fairness and trust in the management, whereas delayed feedback impacts negatively on these variables (Sapienza and Korsgaard, 1996). Receiving timely feedback leads to greater acceptance of decisions (Folger and Konovsky, 1989), as well as their outcomes, even if they are not favorable (Taylor et al., 1992). The term procedural justice refers to the impact of the process of decision making on the quality of exchange relationships (Lind and Tyler, 1988). According to Thibaut and Walker (1975) individuals value fair procedures in particular in situations where direct control over a decision outcome is not possible, because in that case fair procedures provide a means of indirect control. Even in case of a certain negative outcome, just procedures ensure that personal interests are being protected in the long-run. Thus, procedural justice is an important determinant of attitudes and behavior, especially when confronting people with unfavorable outcomes (Lind and Tyler, 1988). Accordingly, procedural justice is positively related to trust in a decision maker (Folger and Konovsky, 1989; Korsgaard et al., 1995), commitment to a decision (Korsgaard et al., 1995), cooperative behavior (Kim and Mauborgne, 1993; Moorman, 1991), and the intention to remain in an existing relationship (Schaubroeck et al., 1994).

Due to the fact that a bulk of studies support the assumption that trust and perceived fairness of authorities are positively linked to tax morale and tax compliance (e.g., Kirchler et al., 2008; Murphy, 2005; Torgler, 2003; Tyler, 2001), it is quite reasonable to assume that long intervals between tax filing and provision of feedback on audits could have a negative impact on perceived procedural fairness and hence – in the long run – on tax honesty. In order to further investigate the relation between delayed feedback, procedural fairness, and tax compliance, the present research investigates the interaction of two different manipulations of procedural fairness. Apart from timing of feedback, in selected periods of taxpaying payoff was reduced by either a decrease in income or an additional compulsory tax. An unheralded collection of an additional tax besides the regular tax due definitely poses a problem with regard to the perception of procedural justice. Thus, the same economic situation, i.e., a lower payoff, is a result of a violation of procedural justice in one condition, but not clearly related to aspects of procedural fairness in another condition.

In a nutshell, the aim of the present research is (1) to replicate the finding that delayed feedback on tax audits results in higher compliance in a different experimental setting, (2) to identify relevant sources of this phenomenon, and (3) to investigate the impact of delayed feedback on perceived fairness, (4) especially in combination with an additional violation of procedural justice. In repeated rounds of taxpaying we test whether a short delay of feedback affects tax compliance. In contrast, Muehlbacher et al. (2012) opted for a one-shot-tax-game with feedback on audits only three weeks after filing taxes in the delayed audit treatment.

As a consequence, one important feature of laboratory experiments on tax compliance applying repeated rounds-designs with feedback in each round has to be considered: the emergence of the so-called “bomb crater effect”. Guala and Mittone (2005), Maciejovsky et al. (2007), as well as Mittone (2006), detected that immediately after an audit, participants’ tax compliance shows a systematic and significant decrease in the subsequent periods. Two different explanations have been offered: (a) Misperception of chance (i.e., participants perceive a lower probability of audit if they were controlled in the previous period) and (b) loss repair (i.e., the tendency to compensate losses due to detected evasion by evading even more in the immediately following rounds). Hence, potential bomb crater effects have to be incorporated when comparing repeated rounds of taxpaying with immediate and delayed feedback on audits. The current study utilizes tax language throughout the instructions and the experiment, with phrases such as “income”, “tax due”, “probability of audit”, and “fine”, which is quite common in scientific research on tax behavior (e.g., Alm et al., 2009; Coricelli et al., 2010; Torgler, 2002), and also claimed to ensure that individuals are aware of the fact that tax honesty is expected (Bruttel and Friehe, 2014), which in turn results in compliance levels closer to real-world findings (Cadsby et al., 2006). Furthermore, participants in our sample were students, but – although sometimes criticized – there is evidence that behavioral responses of student are largely the same as those of nonstudents in the same experiments, as well as that behavioral patterns in the laboratory confirm to those in natural settings (Alm et al., 2015).

## 2. Method

### 2.1. Participants

The present experiment was conducted in the Social Science Research Lab at the Faculty of Psychology at the University of Vienna. Overall, 126 students (76 females and 50 males) with a mean age of 23.07 ( $SD = 3.77$ ) participated. In all experimental conditions participants were students predominantly from social sciences. Participants were not deceived in any way, which is generally not an issue in this respective research lab.

### 2.2. Design and procedure

A repeated rounds design with the dependent variable tax compliance (defined as the average share of tax paid) and the treatment variables feedback (delayed vs. immediate) and payoff reduction (additional compulsory tax vs. reduced income) was implemented. Participants in all 4 conditions (feedback by payoff reduction) had to file taxes in 22 consecutive periods. Their regular income was 3000 ECU (Experimental Currency Units) per round and the tax liability was 900 ECU (corresponding to a tax rate of 30%), and they were informed that both income and tax liability might change over time. The communicated probability of an audit was about 15%, and in case of detection evaded taxes had to be refunded and in addition the same amount had to be paid as a fine. Thus, evading 100 ECU and being caught would result in payment of 200 ECU. All audits were deliberately positioned in periods 3, 12 and 21 in order to keep interfering audit-related effects (e.g., bomb crater effects) constant. Participants were informed that at the end of the session one of the periods would be randomly drawn, and they would be remunerated with respect to the payoff in this respective round according to a fixed exchange rate of 200 ECU = 1 Euro. The relevant period determining the payoff was the same for participants within the same experimental session, but differed between sessions. On average, 11.57 Euros ( $SD = 2.10$ ) were earned.

The basic manipulation referred to the timing of feedback on audits. In the *delayed feedback condition*, participants were informed about performed audits and imposed fines only after the last period (i.e. period 22), whereas in the *immediate feedback condition* this information was given at the end of each period. The additional manipulation reduced the payoff in the periods 8, 9, 10 and 17, 18. In the *compulsory tax condition* the income was 3000 ECU in periods 8, 9, 10 and 17, 18, just as in the other periods, but in addition to the tax due of 900 ECU a compulsory tax of 500 ECU was automatically deducted. In contrast, in the *income reduction condition* participants received only 2500 instead of 3000 ECU income in the respective periods, while the tax due of 900 ECU remained on the level of previous periods. Thus, in both, the compulsory tax condition and the income reduction condition, participants experienced the same decrease in payoff. However, the reason for the reduction of 500 ECU in the respective 5 periods was relevantly different: whereas the compulsory tax was clearly related to perceived fairness of the tax authorities, the income reduction could be attributed to different sources, e.g., ill fortune.

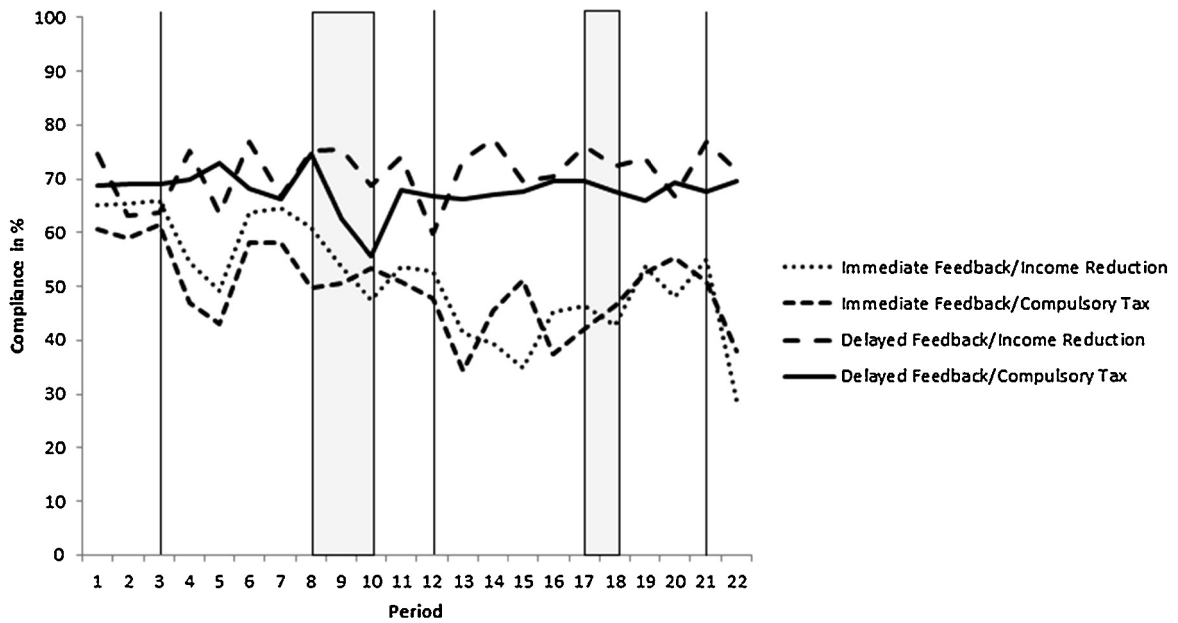
Items of the post-experimental questionnaire addressed the perception of audit probability (“*The probability of detection of tax evasion in the experiment was high.*”), perceived severity of fines (“*When I paid my taxes in the experiment honestly, I did so because the fines for evasion were very severe.*”), perceived fairness with regard to the tax authorities in the experiment (“*The tax authorities in the experiment acted fair toward the taxpayers.*”), and unfairness of the timing of feedback (“*How unfair did you perceive the timing of feedback on whether you were audited?*”). Furthermore, we measured voluntary tax compliance (“*When I paid my taxes in the experiment as required, I did so because to me it is obvious that this is what you do.*”) as well as general tax morale (“*Generally speaking, is cheating on tax never justified, always justified, or something in between?*”). Answers were measured by a Likert-type scale ranging from 1 (do not agree at all) to 9 (totally agree) apart from the general tax morale item, where the scale was from 1 (never justified) to 9 (always justified). In addition, demographic information was recorded. Experimental sessions lasted for about 30 min.

## 3. Results

Fig. 1 illustrates the mean tax compliance in the four experimental conditions over all 22 periods. As can be seen, mean compliance in the delayed feedback conditions is constantly higher than in the immediate feedback conditions, while regarding the payoff reduction manipulation no clear pattern becomes apparent.

Since the dependent variable, i.e. mean tax compliance, was not normally distributed, the effect of the experimental manipulation was tested by tobit regression analysis with the independent variables feedback, payoff reduction, the interaction feedback  $\times$  payoff reduction, and gender. As Table 1 reveals, feedback was a significant predictor of compliance ( $B = -21.00$ ;  $p < .05$ ). The compliance rate was significantly more pronounced in the case of delayed feedback compared to immediate feedback (delayed feedback conditions: 69.49%; immediate feedback conditions: 50.60%). However, neither payoff reduction ( $B = 1.81$ ;  $p = .85$ ) nor the interaction feedback  $\times$  payoff ( $B = 0.46$ ;  $p = .97$ ) were confirmed as significant predictors of compliance. Furthermore, a clear effect of gender was observed ( $B = 19.38$ ;  $p < .01$ ) with females evading less than males (mean compliance females: 65.52%, mean compliance males: 50.92%).

In order to test the assumption that the difference in compliance between the delayed and the immediate feedback conditions develop over time (see Fig. 1), a second tobit regression model was tested including the development over time, i.e. the period, and the interaction feedback  $\times$  period as additional independent variables. As Table 1 shows, the results support this hypothesis, since in this second analysis besides the gender effect ( $B = 40.74$ ;  $p < .01$ ) only the interaction feedback  $\times$  period



**Fig. 1.** Mean compliance over all periods by feedback (delayed vs. immediate) and payoff reduction (additional compulsory tax vs. income reduction) with covariate gender ( $N = 126$ ). The vertical lines indicate audits in period 3, 12 and 21, the rectangles display the periods with reduced wealth (8, 9, 10 and 17, 18).

**Table 1**

Summary of tobit regression analyses with mean relative compliance as dependent variable in both models and feedback, payoff reduction, the interaction feedback  $\times$  payoff reduction and gender as independent variables (model 1;  $N = 126$ ), respectively, feedback, payoff reduction, the interaction feedback  $\times$  payoff reduction, gender, period, and feedback  $\times$  period as independent variables (model 2;  $N = 126$ ).

	Model 1			Model 2		
	<i>B</i>	SE ( <i>B</i> )	<i>p</i>	<i>B</i>	SE ( <i>B</i> )	<i>p</i>
Intercept	57.890	7.871	.000***	59.951	16.402	.000***
Feedback	-20.996	9.180	.024*	-18.045	19.266	.349
Payoff reduction	1.809	9.341	.847	10.773	18.613	.563
Feedback $\times$ payoff reduction	0.463	13.037	.972	0.091	27.377	.997
Gender	19.379	6.689	.004**	40.739	15.114	.007**
Period				0.029	0.426	.624
Feedback $\times$ period				-2.323	0.573	.000***

Note: The variable feedback was dummy coded with 0 = delayed and 1 = immediate feedback; payoff reduction was coded with 0 = additional compulsory tax and 1 = reduced income; gender was coded with 0 = male and 1 = female; period was coded from 1 = period 1 to 22 = period 22.

\*  $p < .05$ .

\*\*  $p < .01$ .

\*\*\*  $p < .001$ .

( $B = -2.32$ ;  $p < .001$ ) was identified as significant predictor of compliance. Feedback ( $B = -18.04$ ;  $p = .35$ ), payoff reduction ( $B = 10.77$ ;  $p = .56$ ), the interaction feedback  $\times$  payoff reduction ( $B = 0.09$ ;  $p = .99$ ) as well as the period ( $B = 0.21$ ;  $p = .62$ ) had no significant influence on the rate of tax compliance. Hence, the difference between delayed and immediate feedback conditions is mainly due to the fact that compliance decreases over time in case of immediate feedback.

To ensure that this strong effect of immediate feedback was not an artifact due to bomb crater effects, we compared the two feedback conditions excluding the respective two rounds directly after the audits identified as bomb craters, i.e., periods 4, 5, 13, 14 and 22.<sup>1</sup> A  $2 \times 2$  ANCOVA with the dependent variable tax compliance and the independent variables feedback (delayed vs. immediate) and pay-off reduction (additional compulsory tax vs. reduced income) as well as the covariate gender confirmed the strong difference in tax compliance between the delayed and the immediate feedback condition ( $F(1, 122) = 7.81$ ,  $p < 0.01$ ; mean delayed feedback = 68.88%; mean immediate feedback = 53.32%).<sup>2</sup>

As stated before, the payoff reduction manipulation did not show an effect on compliance over the course of all 22 periods. Considering exclusively the 5 periods with reduced payoff in an ANCOVA with the dependent variable tax compliance and

<sup>1</sup> Since the experiment ended after period 22, only one round (i.e., period 22) could be excluded after the third and last audit in period 21 in this analysis.

<sup>2</sup> Due to the fact that the dependent variable was not normally distributed we confirmed this difference by a non-parametric test (Mann-Whitney- $U$ :  $z = -2.68$ ,  $p < .01$ ).

**Table 2**

Means and standard deviations (in parentheses) of evaluation of items in the post-experimental questionnaire for the manipulations of feedback and payoff reduction ( $N = 126$ ).

Item	Delayed/compulsory	Delayed/income	Immediate/compulsory	Immediate/income
"The probability of detection of tax evasion in the experiment was high."	4.03 (1.89)	4.34 (2.11)	3.00 (2.07)	3.38 (2.36)
"When I paid my taxes in the experiment honestly, I did so because the fines for evasion were very severe."	5.78 (2.34)	4.86 (2.25)	4.29 (2.56)	3.91 (2.43)
"The tax authorities in the experiment acted fair toward the taxpayers."	4.09 (2.09)	5.41 (2.64)	6.52 (2.38)	6.32 (2.91)
"How unfair did you perceive the timing of feedback on whether you were audited?"	4.66 (1.96)	4.66 (2.45)	3.00 (2.32)	3.15 (2.73)
"When I paid my taxes in the experiment as required, I did so because to me it is obvious that this is what you do."	3.97 (2.42)	5.34 (2.60)	4.39 (2.94)	4.00 (2.86)
"Generally speaking, is cheating on tax never justified, always justified, or something in between?"	3.81 (2.04)	3.59 (1.74)	2.87 (1.77)	2.91 (1.87)

Note: Delayed = delayed feedback; immediate = immediate feedback; compulsory = additional compulsory tax; income = reduced income. Answers were measured by a Likert-type scale from 1 (do not agree at all) to 9 (totally agree), except for the last item (1 – never justified to 9 – always justified).

the independent variables feedback (delayed vs. immediate) and pay-off reduction (additional compulsory tax vs. reduced income) as well as the covariate gender, still no significant difference between the compulsory tax condition and the income reduction condition ( $F(1, 121) = 0.12, p = 0.73$ ; mean compulsory tax = 58.65%; mean income reduction = 60.02%) as well as no interaction feedback  $\times$  payoff reduction could be observed ( $F(1, 121) = 0.02, p = 0.90$ ).<sup>3</sup>

In the post-experimental questionnaire a number of relevant differences concerning the experimental manipulation could be observed: A multivariate analysis with the independent variables delay of feedback and source of payoff reduction revealed that both the perception that the audit probability was high as well as the feeling that fines were severe turned out significantly higher in case of delayed feedback (perception of high audit probability:  $F(1, 122) = 7.31, p < 0.01$ ; mean delayed feedback = 4.18; mean immediate feedback = 3.20; experienced severity of fine:  $F(1, 122) = 8.52, p < 0.01$ ; mean delayed feedback = 5.34; mean immediate feedback = 4.09). In addition, participants in the delayed feedback conditions evaluated the authorities in the experiment as less fair in general ( $F(1, 122) = 13.53, p < 0.001$ ; mean delayed feedback = 4.72; mean immediate feedback = 6.42), and they also rated the timing of feedback as more unfair ( $F(1, 122) = 14.15, p < 0.001$ ; mean delayed feedback = 4.66; mean immediate feedback = 3.08), while there was no significant difference observed for voluntary compliance ( $F(1, 122) = 0.91, p = 0.34$ ; mean delayed feedback = 4.62; mean immediate feedback = 4.18). Regarding the fairness manipulation also a marginally significant interaction of feedback  $\times$  payoff reduction was revealed ( $F(1, 122) = 2.81, p < 0.1$ ): participants experiencing delayed feedback in combination with an additional compulsory tax indicated significantly less fair treatment than all other experimental groups ( $ps < .05$ ). Furthermore, there was a difference with regard to general tax morale, i.e., participants in the delayed feedback condition indicated higher acceptance of the statement that it is ok to cheat on taxes ( $F(1, 122) = 5.96, p < 0.05$ ; mean delayed feedback = 3.71; mean immediate feedback = 2.89). Finally, women showed a higher level of voluntary compliance in comparison to men, independent of the manipulations ( $F(1, 122) = 9.04, p < 0.01$ ; mean women = 4.89; mean men = 3.64). The differences between experimental conditions in the questionnaire are presented in Table 2. Besides the gender effect in voluntary compliance, no other significant differences except those listed in Table 2 were found.

#### 4. Discussion

The present study reveals that delaying feedback on tax audits has a clear effect on tax compliance. Participants receiving immediate feedback on audits and fines evade significantly more of their due tax compared to those who are informed about audits only after the last period. This finding corresponds to the results of Muehlbacher et al. (2012). However, there are some important differences between the two studies, especially with regard to the experimental design. Instead of a single tax decision, the study at hand applied a repeated rounds design. Moreover, the time lag in the delayed feedback condition was considerably shorter, i.e., a few minutes in contrast to three weeks. Most notably, we identify two potential explanations for this prominent impact of timing of feedback on tax compliance: Although informed about objective probability of audit as well as the precise level of fines, people in the delayed feedback group were more likely to state that the probability of audit

<sup>3</sup> The finding that there was no difference between the reduced payoff conditions was also confirmed by a non-parametric test (Mann–Whitney–U:  $z = -0.46, p < .65$ ).

was high, and they perceived fines to be more severe than participants with immediate feedback. This outcome corresponds to reports that rare events are weighted differently depending on the fact if only descriptions of probabilities are available or if they are directly experienced (Hertwig et al., 2004; Hertwig and Erev, 2009). Hence, delayed feedback seems to promote overweighting of audit probability and thus a higher level of compliance. The finding that the difference in compliance develops over time supports the assumption that experienced audit probability is a crucial factor for the emergence of this notable difference in compliance. Regarding the payoff reduction either due to an additional compulsory tax or by reduced income, no different impact on tax compliance was observed. One reason for this might be that the number of periods with reduced wealth was too few to trigger an observable reaction. Additionally, a difference between a compulsory tax and less income should be given in reality, but is hard to manipulate in a rather artificial laboratory setting, since both reductions may be attributed to the experimenter in the end. Altogether, the behavioral results suggest that a longer time interval between filing taxes and learning about the occurrence of an audit is a useful tool to increase tax honesty.

In contrast, the analysis of fairness-related measures can be interpreted in a different way. Participants with delayed feedback perceive the timing of information on audits as significantly more unfair, and furthermore evaluate the respective authorities as acting unfairly. This finding corresponds to the literature claiming that feedback in general has to be provided in time in order to shape behavior in a positive way (Sapienza and Korsgaard, 1996). In this vein it may be shortsighted to reason that honest taxpayers comply independent from whether they get feedback on the consequences of an audit immediately or with a delay. In line with the established connection of perceived fairness, trust in authorities, and tax compliance (e.g., Bergmann, 2002; Fjeldstad, 2004; Kirchler et al., 2014; Pommerehne and Frey, 1992), it seems plausible that the perception of unfair treatment could prompt honest taxpayers to abandon compliance. However, here it is relevant to differentiate between delayed audits with feedback subsequently and audits where the consequences are not communicated within a substantial time lag, because it is quite reasonable to assume that the latter case might be perceived as more unfair. Since participants in our experiment might have interpreted the experimental manipulation as a delay of audit or a delay of feedback on an audit already completed, this might have influenced fairness evaluations but cannot be clearly disentangled in the analyses.

Concerning the somehow inconsistent effect of delayed feedback on compliance and perceived fairness, one possible interpretation could be that long periods of uncertainty may yield higher honesty for the short term, but nevertheless undermine compliance in the long run. In particular, when individuals are unable to have direct control in a situation, they are more apt to pay attention to the relevant procedures (Leung and Li, 1990). However, the present experiment does not offer any further insight on that since we did not investigate any long-term effects. For instance, the behavior of participants getting delayed information on the occurrence of audits could change considerably after getting feedback for the first time. Thus, reactance provoked by delayed feedback might result in reluctance to comply with the tax laws after the first response, presumably enhancing strategic decisions afterwards. This conclusion is also maintained by the fact that regarding tax morale in general, participants facing delayed feedback show a significantly higher acceptance of tax cheating in real life situations. Besides that, informing participants only at the end of the experiment about the period randomly selected to determine their payoff might also have evoked negative feelings or reactance as a form of violation of procedural fairness. However, this specific feature of our experiment was of course present in all experimental conditions, so it cannot account for the differences between treatments.

The result that women showed higher compliance in the experiment as well as a higher intended voluntary compliance in the analysis of the questionnaire is in line with numerous other findings in the literature (e.g., Hasseldine, 1999; Kastlunger et al., 2010; Kogler et al., 2013). Women seem to be more honest taxpayers in experiments as well as in reality, showing a stronger motivation to comply in absence of enforcement mechanisms.

Finally, the current findings bear important implications for experimental tax research: settings introducing delayed feedback on audits might represent essential features of the real taxpaying situation better than the majority of tax experiments providing immediate feedback on audits. In the best case, a different setting just produces a shift in the general level of compliance, but at least in combination with specific additional treatments, unintended interaction effects might lead to biased and not generally valid results.

## Acknowledgements

We thank the editors, two anonymous reviewers and Jerome Olsen from the University of Vienna for their helpful comments.

## References

- Alm, J., Jackson, B.R., McKee, M., 2009. *Getting the word out: enforcement information dissemination and compliance behavior*. *J. Public Econ.* 93, 392–402.
- Alm, J., Bloomquist, K.M., McKee, M., 2015. *On the external validity of laboratory tax compliance experiments*. *Econ. Inq.* 53 (2), 1170–1186.
- Badia, P., McBane, B., Suter, S., Lewis, P., 1966. *Preference behavior in an immediate versus variably delayed shock situation with and without a warning signal*. *J. Exp. Psychol.* 72, 847–852.
- Bergmann, M., 2002. *Who pays for social policy? A study on taxes and trust*. *J. Soc. Policy* 31, 289–305.
- Bruttel, L., Friehe, T., 2014. *On the path dependence of tax compliance*. *Eur. Econ. Rev.* 65, 90–107.
- Cadsby, C.B., Maynes, E., Trivedi, V., 2006. *Tax compliance and obedience to authority at home and in the lab: a new experimental approach*. *Exp. Econ.* 9 (4), 343–359.
- Coricelli, G., Joffily, M., Montmarquette, C., Villeval, M.C., 2010. *Cheating, emotions, and rationality: an experiment on tax evasion*. *Exp. Econ.* 13 (2), 226–247.

- Fjeldstad, O.-H., 2004. What's trust got to do with it? Non-payment of service charges in local authorities in South Africa. *J. Mod. Afr. Stud.* 42, 539–562.
- Folger, R., Konovsky, M.K., 1989. Effects of procedural and distributive justice on reactions to pay raise decisions. *Acad. Manag. J.* 32, 115–130.
- Guala, F., Mittone, L., 2005. Experiments in economics: external validity and the robustness of phenomena. *J. Econ. Methodol.* 12, 495–515.
- Hasseldine, J., 1999. Gender differences in tax compliance. *Asia-Pac. J. Tax.* 3, 73–89.
- Hertwig, R., Barron, G., Weber, E.U., Erev, I., 2004. Decisions from experience and the effect of rare events in risky choice. *Psychol. Sci.* 15 (8), 534–539.
- Hertwig, R., Erev, I., 2009. The description–experience gap in risky choice. *Trends Cogn. Sci.* 13 (12), 517–523.
- Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. *Econometrica* 47, 263–291.
- Kastlunger, B., Dressler, S., Kirchler, E., Mittone, L., Voracek, M., 2010. Sex differences in tax compliance: differentiating between demographic sex, gender-role orientation, and prenatal masculinization (2D:4D). *J. Econ. Psychol.* 31, 542–552.
- Kim, W.C., Mauborgne, R.A., 1993. Procedural justice, attitudes, and subsidiary top management compliance. *Acad. Manag. J.* 36, 502–526.
- Kirchler, E., Hoelzl, E., Wahl, I., 2008. Enforced versus voluntary tax compliance: the slippery slope framework. *J. Econ. Psychol.* 29 (2), 210–225.
- Kirchler, E., Kogler, C., Muehlbacher, S., 2014. Cooperative tax compliance: from deterrence to deference. *Curr. Dir. Psychol. Sci.* 23, 87–92.
- Kogler, C., Batrancea, L., Nichita, A., Pantya, J., Belianin, A., Kirchler, E., 2013. Trust and power as determinants of tax compliance: testing the assumptions of the slippery slope framework in Austria, Hungary, Romania and Russia. *J. Econ. Psychol.* 34, 169–180.
- Korsgaard, M.A., Schweiger, D.M., Sapienza, H.J., 1995. The role of procedural justice in building commitment, attachment, and trust in strategic decision-making teams. *Acad. Manag. J.* 38, 60–84.
- Loewenstein, G., 1987. Anticipation and the valuation of delayed consumption. *Econ. J.* 97, 666–684.
- Leung, K., Li, W., 1990. Psychological mechanisms of process–control effects. *J. Appl. Psychol.* 75, 613–620.
- Lind, E.A., Tyler, T., 1988. *The Social Psychology of Procedural Justice*. Plenum, New York.
- Maciejovsky, B., Kirchler, E., Schwarzenberger, H., 2007. Misperception of chance and loss repair: on the dynamics of tax compliance. *J. Econ. Psychol.* 28, 678–691.
- Mittone, L., 2006. Dynamic behaviour in tax evasion: an experimental approach. *J. Socio-Econ.* 35, 813–835.
- Moorman, R.H., 1991. Relationship between organizational justice and organizational citizenship behaviors: do fairness perception influence employee citizenship? *J. Appl. Psychol.* 76, 845–855.
- Muehlbacher, S., Mittone, L., Kastlunger, B., Kirchler, E., 2012. Uncertainty resolution in tax experiments: why waiting for an audit increases compliance. *J. Socio-Econ.* 41, 289–291.
- Murphy, K., 2005. Regulating more effectively: the relationship between procedural justice, legitimacy, and tax non-compliance. *J. Law Soc.* 32, 562–589.
- Pommerehne, W., Frey, B.S., 1992. *The Effects of Tax Administration on Tax Morale (Series II, No. 191)*. University of Konstanz.
- Sapienza, H.J., Korsgaard, M.A., 1996. Procedural justice in entrepreneur–investor relations. *Acad. Manag. J.* 39, 544–574.
- Schaubroeck, J., May, D.R., Brown, F.W., 1994. Procedural justice explanations and employee reactions to economic hardship: a field experiment. *J. Appl. Psychol.* 79, 455–460.
- Taylor, M.S., Tracy, K.R., Harrison, J.K., Renard, M.K., Carroll, S.J., 1992. Testing the due process metaphor for performance appraisal systems: a filed application. In: *Paper Presented at the 50th Annual Meeting of the Academy of Management, Las Vegas*.
- Thibaut, J., Walker, L., 1975. *Procedural Justice: A Psychological Analysis*. Erlbaum, Hillsdale, NY.
- Torgler, B., 2002. Speaking to theorists and searching for facts: tax morale and tax compliance in experiments. *J. Econ. Surv.* 16 (5), 657–683.
- Torgler, B., 2003. Tax morale, rule-governed behaviour and trust. *Const. Polit. Econ.* 14, 119–140.
- Tyler, T.R., 2001. Public trust and confidence in legal authorities: what do majority and minority group members want from the law and legal institutions? *Behav. Sci. Law* 19, 215–235.
- Zeigarnik, B., 1927. Das Behalten erledigter und unerledigter Handlungen. *Psychologische Forschungen* 9, 1–85.