Social Cognitive and Affective Neuroscience, 2017, 1108–1117

doi: 10.1093/scan/nsx029
Advance Access Publication Date: 11 April 2017
Original article

Coercive and legitimate authority impact tax honesty: evidence from behavioral and ERP experiments

Katharina Gangl,[1,2,3] Daniela M. Pfabigan,[4,5] Claus Lamm,[4] Erich Kirchler,[1] and Eva Hofmann[1,6]

[1]Department of Applied Psychology: Work, Education, Economy; Faculty of Psychology, University of Vienna, Vienna, Austria, [2]Economic and Social Psychology, Institute of Psychology, University of Goettingen, Goettingen, Germany, [3]Economic Psychology and Leadership Ethics Unit, Leadership Excellence Institute Zeppelin, Zeppelin University, Friedrichshafen, Germany, [4]Department of Basic Psychological Research and Research Methods, Social, Cognitive and Affective Neuroscience Unit, Faculty of Psychology, University of Vienna, Vienna, Austria, [5]Culture and Social Cognitive Neuroscience Laboratory, School of Psychological and Cognitive Sciences, Peking University, Beijing, China, and [6]Centre for Trust, Peace and Social Relations, Coventry University, Coventry, United Kingdom

Katharina Gangl and Daniela M. Pfabigan contributed equally to this work.
Correspondence should be addressed to Katharina Gangl, University of Goettingen, Institute of Psychology, Economic and Social Psychology Unit, Goßlerstraße 14, 37073 Göttingen, Germany.
E-mail: gangl@psych.uni-goettingen.de.

Abstract

Cooperation in social systems such as tax honesty is of central importance in our modern societies. However, we know little about cognitive and neural processes driving decisions to evade or pay taxes. This study focuses on the impact of perceived tax authority and examines the mental chronometry mirrored in ERP data allowing a deeper understanding about why humans cooperate in tax systems. We experimentally manipulated coercive and legitimate authority and studied its impact on cooperation and underlying cognitive (experiment 1, 2) and neuronal (experiment 2) processes. Experiment 1 showed that in a condition of coercive authority, tax payments are lower, decisions are faster and participants report more rational reasoning and enforced compliance, however, less voluntary cooperation than in a condition of legitimate authority. Experiment 2 confirmed most results, but did not find a difference in payments or self-reported rational reasoning. Moreover, legitimate authority led to heightened cognitive control (expressed by increased MFN amplitudes) and disrupted attention processing (expressed by decreased P300 amplitudes) compared to coercive authority. To conclude, the neuronal data surprisingly revealed that legitimate authority may led to higher decision conflict and thus to higher cognitive demands in tax decisions than coercive authority.

Key words: taxpaying; cognitive control; social decision-making; MFN; P300; P2

Received: 22 June 2016; **Revised:** 14 February 2017; **Accepted:** 1 March 2017

Introduction

Social systems are characterized by individuals who cooperate in the provision of non-excludable public goods, or defect and free-ride (Dawes, 1980). Tax honesty and tax evasion represent highly relevant examples of such cooperation. They do not only impact the possibilities of modern societies to provide schools or roads, but can also fuel social turbulences based on perceived injustice, if some taxpayers are perceived to systematically avoid paying their fair share of taxes. To uphold cooperation in these fragile systems, modern societies employ centralized institutions who wield *harsh* coercive or *soft* legitimate authority to ensure individual cooperation (Raven *et al.*, 1998; Turner, 2005; Güth *et al.*, 2007; Baldassarri and Grossman, 2011). Empirical evidence shows that the two qualities of authority lead to different motivations to cooperate (Kastlunger *et al.*, 2012; Hofmann *et al.*, 2014; Hofmann *et al.*, 2017), but rarely differ in their positive effect on cooperation (Alm *et al.*, 1992; Tyler and Fagan, 2008; Isakov and Rand, 2012; Hofmann *et al.*, 2014; Hartl *et al.*, 2015).

Coercive and legitimate authority promote cooperation originating from different motivational states (Koslowsky *et al.*, 2001; Kirchler *et al.*, 2008; Gangl *et al.*, 2015). Coercive authority based on control and punishment is assumed to fuel mistrust (Mulder *et al.*, 2006), to be perceived as ferocious and unfair (Mossholder *et al.*, 2009), and to induce calculative cost-benefit analyses (Kirchler, 2007), which frame the decision to cooperate as a business decision rather than as an ethical decision (Tenbrunsel and Messick, 1999). Thus, the decision to cooperate might become a deliberate rational decision about one's own egoistic utility and societal goals become less salient. Consequently, coercive authority is seen to cause resistance, reactance (Brehm, 1966) and negative emotional arousal (Coricelli *et al.*, 2010), and it enforces motivation to comply out of feared penalization (Frey, 1997; Kirchler *et al.*, 2008). In contrast, legitimate authority is based on acceptance, appreciation, perceived expertise and information provision (Raven *et al.*, 1998; Gangl *et al.*, 2015). It is assumed to increase trust and fairness (Tyler and Fagan, 2008; Hechter, 2009) and might be perceived as the communication of a widely accepted rule by community members (Bendor and Mookherjee, 1990), or as an instrument to establish fairness and a social norm of cooperation (Fehr and Fischbacher, 2002). Legitimate authority might induce moral pressure and psychological stress (e.g. a conflict, Erard and Feinstein, 1994; Dulleck *et al.*, 2016) in response to the felt urge to reciprocate to legitimate authority (Fehr *et al.*, 1997). Legitimate authority might provoke a deliberate conflict between self-interest and community interests which, in turn, fosters a voluntary motivation to cooperate (Tyler, 1997; Kirchler *et al.*, 2008; Hofmann *et al.*, 2014). Thus, decisions about cooperation under coercive or legitimate authority do reflect rational economic choices (Camerer, 2003) *and* social motives of reciprocity or inequity aversion (Fehr and Fischbacher, 2002). Although coercive and legitimate authority are perceived so differently, their positive effects on cooperative behavior do not seem to vary dramatically (Hartl *et al.*, 2015).

Taxpaying can be considered as a highly relevant real life example for cooperation in a social context. Tax behavior is frequently investigated in laboratories. Referring to experimental games, income, tax rates, audit probabilities and fine rates are manipulated and tax payments over several tax filing trials are registered as dependent variable (Mittone, 2006; Alm *et al.*, 2010; Hartl *et al.*, 2015). Although a bulk of empirical studies on tax behavior has been published, studies on cognitive processes underlying tax behavior are rare and neurophysiological insights are missing.

We performed an event-related potentials (ERPs) study, investigating how tax payment decisions are affected by different types of centralized authorities. ERPs have been repeatedly used to gain insights into social decision-making, with their high temporal resolution enabling detailed insights into the mental chronometry of decision-making. In particular, amplitude variation of the Medial Frontal Negativity component (MFN; Gehring and Willoughby, 2002) has been investigated in this regard. It is a negative scalp potential within 200–300 ms after (feedback) stimulus onset at fronto-central electrodes, with the anterior midcingulate cortex (aMCC) and ventral striatum as potential neuronal generators (Debener *et al.*, 2005; Becker *et al.*, 2014). Generally, MFN amplitudes are more negative after unfavorable compared to favorable (Miltner *et al.*, 1997), unexpected compared to expected (Alexander and Brown, 2011; Pfabigan *et al.*, 2011), and salient compared to insignificant outcomes (Talmi *et al.*, 2013). Moreover, MFN enhancement was interpreted as a signal of enhanced cognitive control (Botvinick *et al.*, 2001; Yeung *et al.*, 2004), which describes, among others, the monitoring and regulation of response strategies, feedback processing yielding strategy regulation and response conflict (i.e. concurrent activation of incompatible response options; Folstein and Van Petten, 2008). Taken together, MFN amplitude variation reflects early and coarse stimulus evaluation processes (Yeung *et al.*, 2004; Hajcak *et al.*, 2006). Regarding social decision-making, the MFN is usually observed in economic games focusing on asset distribution, such as the Ultimatum Game (UG; Güth *et al.*, 1982). There, MFN amplitude enhancement indicates unfair compared to fair outcomes (Boksem and De Cremer, 2010; Alexander and Brown, 2011; Alexopoulos *et al.*, 2012, 2013), concurrently also reflecting cognitive control.

Another ERP component frequently investigated during social decision-making is the P300. This positive-going component peaks within 300–500 ms after (feedback) stimulus onset at posterior electrodes (Polich, 2007). More pronounced P300 amplitudes have been reported in response to a greater change in evaluative stimulus categorization (Cacioppo *et al.*, 1993; Ito *et al.*, 1998) and greater attention allocation (Polich, 2007). Larger P300 amplitudes were found after positive compared to negative outcomes (Bellebaum and Daum, 2008; Pfabigan *et al.*, 2015). In UG studies, P300 amplitudes were also sensitive to outcome fairness–usually larger after fair compared to unfair offers (Wu *et al.*, 2012; Qu *et al.*, 2013).

We conducted two experiments introducing taxpaying in two centralized institutional contexts. In the first experiment, we collected behavioral data and reaction times while participants performed repeated fast tax decisions under coercive and legitimate authority. In the second experiment, the tax paradigm was slightly adapted to an individual assessment and such that the speed of participants' proceeding the tax decision was reduced, to allow a more valid assessment of ERP data.

As in previous studies, we expected that in contrast to legitimate authority, coercive authority will lead to self-reported enforced motivation to comply, low voluntary motivation to cooperate (Kastlunger *et al.*, 2012; Hofmann *et al.*, 2014) and enhanced rational decision-making and reactance than legitimate authority (Brehm, 1966, Kirchler, 2007). Overall tax payments should not differ between the authority conditions (Hartl *et al.*, 2015). For reaction times (experiment 1) and ERP data (experiment 2) we tested two opposing hypotheses. On the one hand, coercive authority might induce a more complex decision process than legitimate authority. It provokes a comparatively

effortful calculative cost-benefit analysis and deliberation about possible ways to reduce the tax burden. Legitimate authority might be perceived as a positive signal and thus produces a fast response to spontaneously cooperate (Tenbrunsel and Messick, 1999; Kirchler, 2007). On the other hand, legitimate authority might induce more complex cognitive processes than coercive authority, as personal utility optimization motives might be in conflict with the moral pressure to pursue societal goals. Reciprocity towards the legitimate authority might elicit a cognitive response conflict whereas coercive authority produces a fast response to cooperate (Fehr *et al.*, 1997). Reaction times serve as indicator of cognitive effort and indicate whether cost-benefit analysis induced by coercive authority or the heightened response conflict induced by legitimate authority lead to more complex, deliberate and, thus, slower decisions (Rubinstein, 2007). For the ERP data, we expected MFN enhancement to indicate heightened cognitive control demands due to response conflict either in the coercive or the legitimate context (Botvinick *et al.*, 2001). Enhanced P300 amplitudes should reflect enhanced categorization demands and attentional processing (Cacioppo *et al.*, 1993; Ito *et al.*, 1998; Polich, 2007) in case either coercive or legitimate authority induce a more complex decision process. In addition, we explored the P2 component prior to the MFN, which is indicative of arousal levels (Carretié *et al.*, 2001) and attention capture (Potts, 2004).

Experiment 1

Method

The following section describes the central methodological aspects of our study; additional methodological details are provided in Supplementary Material.

Sample

The sample consisted of 80 volunteers (38 men, 1 did not indicate sex; $M_{age} = 24.89$, $SD_{age} = 6.48$). Participants were randomly assigned to one of two taxpaying conditions (coercive authority of tax administration followed by legitimate authority [$n = 39$], or legitimate authority followed by coercive authority, [$n = 41$]. The study was conducted in accordance with the Declaration of Helsinki (7th revision, 2013) and local ethical guidelines for experimentation with human participants (including approval by an institutional review board) at the Faculty of Psychology, University of Vienna.

Procedure

Participants were asked to imagine being self-employed, earning money in Experimental Currency Units (ECU) and paying taxes (40% of their income) over several trials in a fictitious country. They were informed that the tax administration would conduct tax audits with a chance of 15%. In case of detected tax evasion they would have to pay back the evaded amount plus a fine of the same amount. Participants were informed that a randomly chosen trial determinates their final remuneration for participation. After the introduction to the rules of the tax game, in counterbalanced order, 40 coercive and 40 legitimate authority trials followed, presented on a PC. After introducing the first authority, each trial started with the endowment of one out of five incomes (each presented eight times per authority in exactly the same sequence for all participants; the randomized sequence was determined prior to the experiment) and the presentation of 40% tax rate. Then participants decided how much tax they wanted to pay. After 40 trials in the first country, participants filled out the first self-report questionnaire. Subsequently, they were told to move to another country with a changed authority where they filed taxes during another 40 trials, ending by filling out the second self-report questionnaire. Participants were reminded of the respective authority manipulation every 10th trial. Finally, one prior to the experiment randomly selected trial determined the payment of all participants based on their behavior in this trial. On average, individuals earned €10.78.

Material

To manipulate authority, two scenarios were developed which described tax administrations that basically work with harsh controls and punishments (coercive authority) or with professional experts who support taxpayers filing their taxes (legitimate authority). To allow a direct comparison of the effect of both scenarios, a within-subjects design was used in which all participants were presented with both scenarios, in counterbalanced order. A manipulation check revealed that the centralized institutions were perceived as manipulated (see Supplementary Material [1.1.3, 3] for material and details on the manipulation check).

Tax behavior was assessed by averaging taxpaying decisions over 40 tax filing trials. For each tax filing trial one out of five incomes (60.000 ECU, 70.000 ECU, 75.000 ECU, 85.000 ECU, or 90.000 ECU) was assigned to the participants. Participants read the information about coercive or legitimate authority, their income, and the tax rate until they choose to terminate the slide by pressing key 1. Afterwards, participants were presented with a slide with five possible tax amounts to pay corresponding to 0% (fully dishonest), 25%, 50%, 75% and 100% (fully honest) of the tax due. This slide was presented until participants choose one of the options via button press (keys 1–5, pre-assigned to the respective percent value). The different income levels ensured that participants varied the responses over the 40 trials.

Mean reaction times were assessed from the onset of the tax amount screen until button press. For each trial, outliers (2.17% of all data points) were discarded in case reaction times were longer than the mean reaction time plus three times the standard deviation.

Self-reports were assessed with two identical questionnaires. Perception of coercive authority and legitimate authority was assessed with the Interpersonal Authority Inventory (Raven *et al.*, 1998; Hofmann *et al.*, 2014) adapted to our tax context. Reactance was measured by the adapted Hong Psychological Reactance Scale (Shen and Dillard, 2005). Deliberate rational decision-making was measured by the adapted Rational-Experiential Inventory (Pacini and Epstein, 1999). Enforced compliance was measured by the TAX-I (Kirchler and Wahl, 2010), voluntary cooperation by items adapted from the motivational posture scale capitulation (Braithwaite, 2003). Response scales were 7-point Likert scales (1–disagree, 7–agree). Items and their reliability scores are reported in Supplementary Material (4).

Statistical analyses

To analyze the differential impact of the two centralized institutions, 2 x 2 univariate analysis of variance with authority manipulation (coercive *vs* legitimate) as within-subject factors, and order of manipulation (coercive followed by legitimate authority

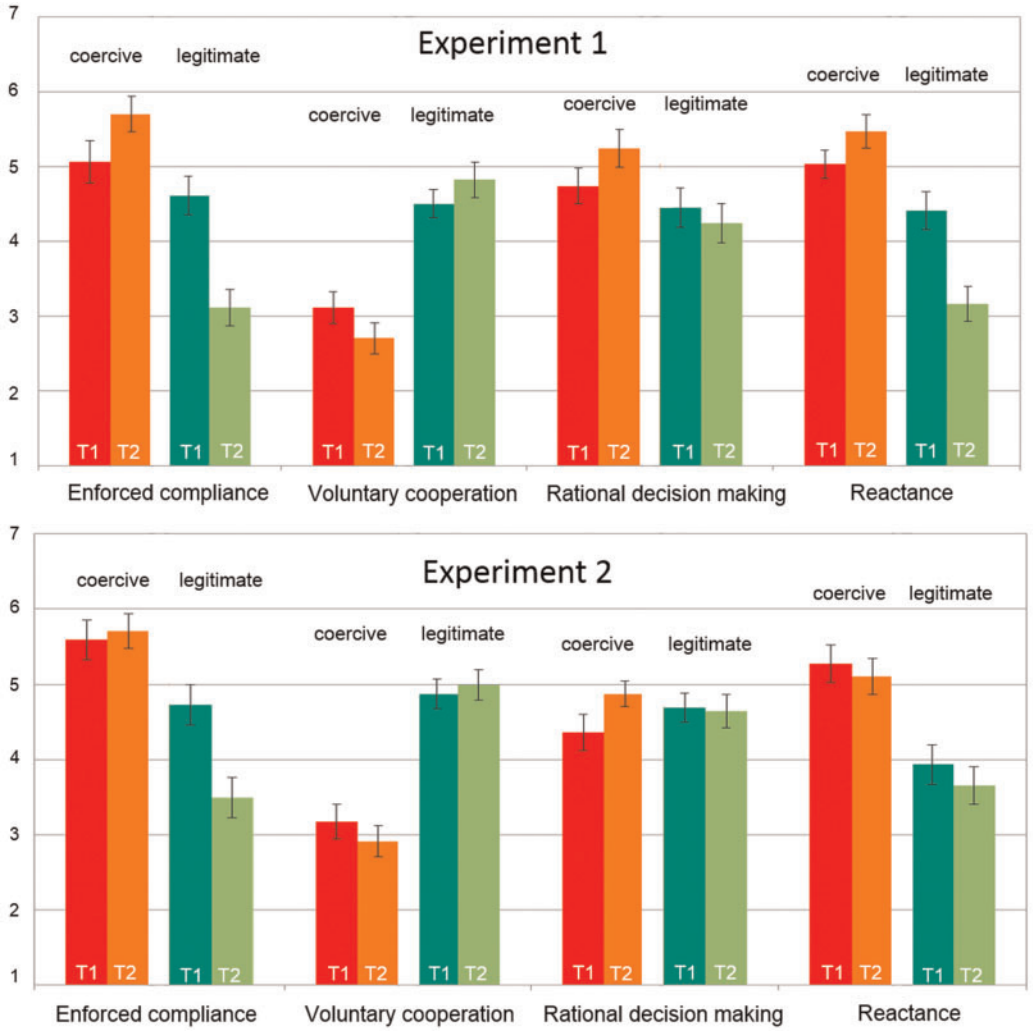

Fig. 1. Differences in self-reports in experiment 1 and 2.
Note: The bars represent means with the standard error. T1 and T2 indicate whether the authority was presented first (T1, i.e. first 40 trials) or second (T2, i.e. second 40 trials). 1 = disagree, 7 = agree.

or vice versa) as between-subjects factor was performed, with self-report, behavioral and reaction time data as dependent variables. Effect sizes of significant results are specified with partial eta squared (η_p^2) and *Cohen's d*; the alpha-level was set at $P < 0.05$. In the following, we present only the significant results of main effects. Detailed results and robustness checks are presented in Supplementary Material (1.1.4, 1.1.5).

Results

Self-report data. Self-reported enforced compliance was significantly affected by the type of authority ($F(1,77) = 54.55$, $P < 0.001$, $\eta_p^2 = 0.42$). As Figure 1 shows, enforced compliance was higher in the coercive ($M = 5.39$, $SE = 0.18$) than in the legitimate authority condition ($M = 3.88$, $SE = 0.20$). This is especially true if coercive authority follows legitimate (for details on an interaction effect see Supplementary Material). On the other hand, voluntary cooperation was only affected by authority ($F(1,77) = 59.51$, $P < 0.001$, $\eta_p^2 = 0.44$), it was higher under legitimate ($M = 4.65$, $SE = 0.15$) than coercive authority ($M = 2.90$, $SE = 0.15$). Also, rational decision-making was affected by authority ($F(1,78) = 11.55$, $P < 0.001$, $\eta_p^2 = 0.13$) and was higher under coercive ($M = 5.00$, $SE = 0.17$) than legitimate authority ($M = 4.35$, $SE = 0.18$). Reactance was also higher under coercive ($M = 5.25$, $SE = 0.15$) than legitimate authority ($M = 3.80$, $SE = 0.18$; $F(1,78) = 81.77$, $P < 0.001$, $\eta_p^2 = 0.51$; for details see Supplementary Material).

Behavioral data. Tax compliance was significantly higher under legitimate ($F(1,78) = 5.96$, $P = 0.017$, $\eta_p^2 = 0.07$; $M = 4.07$; $SE = 0.12$) than coercive authority ($M = 3.77$, $SE = 0.14$). Reaction times were significantly affected by the contrast between coercive and legitimate authority ($F(1,78) = 177.63$, $P < 0.001$, $\eta_p^2 = 0.70$). Compared to time 1 (first 40 trials) ($t(78) = 1.05$, $P = 0.295$), at time 2 (second 40 trials) ($t(78) = -1.97$, $P = 0.052$, $d = 0.45$) coercive authority ($M = 1784.67$ ms, $SE = 138.61$) by trend yielded faster responses than legitimate authority ($M = 2230.02$ ms, $SE = 138.98$).

Discussion experiment 1

Corroborating previous studies, experiment 1 shows that a perceived coercive compared to a legitimate authority led to less

voluntary cooperation (Hofmann *et al.*, 2014). For enforced compliance and reactance, we observed sensitivity to the change from coercive to legitimate authority. More specifically, legitimate authority's negative effect on enforced compliance and reactance was particularly strong when a legitimate authority was perceived as the direct change following coercive authority. Overall and in line with previous studies (e.g. Hofmann *et al.*, 2014), the effect sizes indicated that the difference between coercive and legitimate authority was the most important factor affecting enforced compliance and reactance. Results also showed that legitimate authority led to slightly higher tax payments than coercive authority. Coercive authority induced more self-reported rational decision-making than legitimate authority. However, reaction time results indicate that the decision under coercive authority was quicker to reach compared to legitimate authority. At time 2 (second 40 trials), coercive authority led to faster reactions than legitimate authority. Experiment 1 is one of the rare attempts showing that the different psychological meanings of coercive and legitimate authority can also be detected by indirect measures of cognitive processes such as reaction times. To clarify the exact nature of these processes and how coercive and legitimate authority affect them, data providing access to the temporal sequence and different cognitive processes underlying decision-making and behavior are needed. Therefore, our second experiment assessed ERPs while participants made tax decisions under different tax authorities.

Experiment 2

Method

Sample

The sample consisted of 81 volunteers of which three were excluded because of limited task comprehension or excessive alpha-band EEG activity (final sample: $n = 78$, 40 men, $M_{age} = 24.51$, $SD_{age} = 5.05$). All participants were right-handed (Oldfield, 1971), had normal or corrected-to-normal vision, and reported no past or present neurological or psychiatric disorder. All participants gave written informed consent prior to the experiment. Recruitment and ethical guidelines were comparable to experiment 1. They were again randomly assigned to one of two taxpaying conditions (coercive authority of tax administration followed by legitimate authority [$n = 38$], or legitimate authority followed by coercive authority [$n = 40$]).

Procedure

Participants were tested individually in a sound-attenuated, shielded chamber. As in experiment 1, they were asked to imagine being self-employed, earning income, and paying taxes over several years. To ensure that participants understood the instructions, they were led through a taxpaying example and performed 10 training trials. Then, identical to experiment 1 (order counterbalanced), 40 coercive and 40 legitimate authority trials followed. However, different to experiment 1, participants were reminded every 5th trial to make the manipulation more salient. Each trial started with the presentation of a white fixation cross on black background (1000ms). Afterwards, the flag of the respective country, the fictitious income, and the 40% tax rate in total numbers were presented centrally on black background (3000 ms). Prior to the actual tax decision, another screen was blended in asking 'How much tax would you pay in [current country]?' (2000 ms). This slide was added to the experimental design to avoid fast responses to the tax decision as in Experiment 1. Afterwards, participants were presented with five options for taxpaying, either representing 0%, 25%, 50%, 75% or 100% of the 40% tax rate (absolute numbers). Participants had to choose between the keys 1–5 on a standard keyboard to indicate their tax decision via button press. No time limit was given for their decision. A variable inter-stimulus-interval (1400–1600 ms) was presented afterwards depicting the fixation cross. After 40 trials in the first context, participants filled out the first self-report questionnaire and were subsequently informed that they would move to a different country with a different tax administration. After introducing the second context, another 40 trials were presented applying the same experimental design as before, ending with filling out the second self-report questionnaire. At the end, participants were remunerated based on a €10 show-up fee and their tax decision in one randomly chosen tax trial (on average €22.12).

Material

The same material as in experiment 1 was used (Supplementary Material 1.2.3).

EEG was recorded from 57 equidistantly arranged electrodes in a cap. Signal preprocessing and artefact correction was conducted using EEGLAB (Delorme and Makeig, 2004). Please refer to Supplementary Material for further details on data collection and preprocessing (1.2.2).

To assess ERP amplitudes, data were epoched time-locked to the onset of the income separately for coercive and legitimate trials. MFN amplitudes were extracted at FCz as peak-to-peak-to-peak values (Yeung and Sanfey, 2004); i.e. difference between the MFN component and the mean of the preceding P2 and the subsequent P300 component) in the time window 150–400 ms post income. P300 amplitudes were extracted at Pz as peak-to-peak values (Pfabigan *et al.*, 2011); i.e. difference between the P300 and the preceding N2 component in the time window 200–600 ms post income.

Statistical analyses

Several 2 x 2 univariate analyses of variance with manipulation of authority (coercive *vs* legitimate authority) as within-subjects factors, and order of manipulation (coercive followed by legitimate authority and vice versa) as between-subjects factor were performed with self-report data, behavioral data, and P2, MFN and P300 amplitudes as dependent variables. Again, we present only significant results. Detailed results and robustness checks are presented in Supplementary Material (1.2.4–6).

Results

Self-report data. As shown in Figure 1, self-reported enforced compliance was again significantly higher under coercive ($M = 5.65$, $SE = 0.17$) than legitimate authority ($M = 4.13$, $SE = 0.20$; $F(1,76) = 50.78$, $P < 0.001$, $\eta_p^2 = 0.40$). Again, voluntary cooperation was significantly higher under legitimate ($M = 4.92$, $SE = 0.14$) than coercive authority ($M = 3.03$, $SE = 0.15$; $F(1,76) = 103.11$, $P < 0.001$, $\eta_p^2 = 0.58$). Rational decision-making did not differ between coercive and legitimate authority. Again, reactance was significantly higher under coercive ($M = 5.18$, $SE = 0.17$) than legitimate authority ($M = 3.79$, $SE = 0.18$; $F(1,75) = 50.17$, $P < 0.001$, $\eta_p^2 = 0.40$).

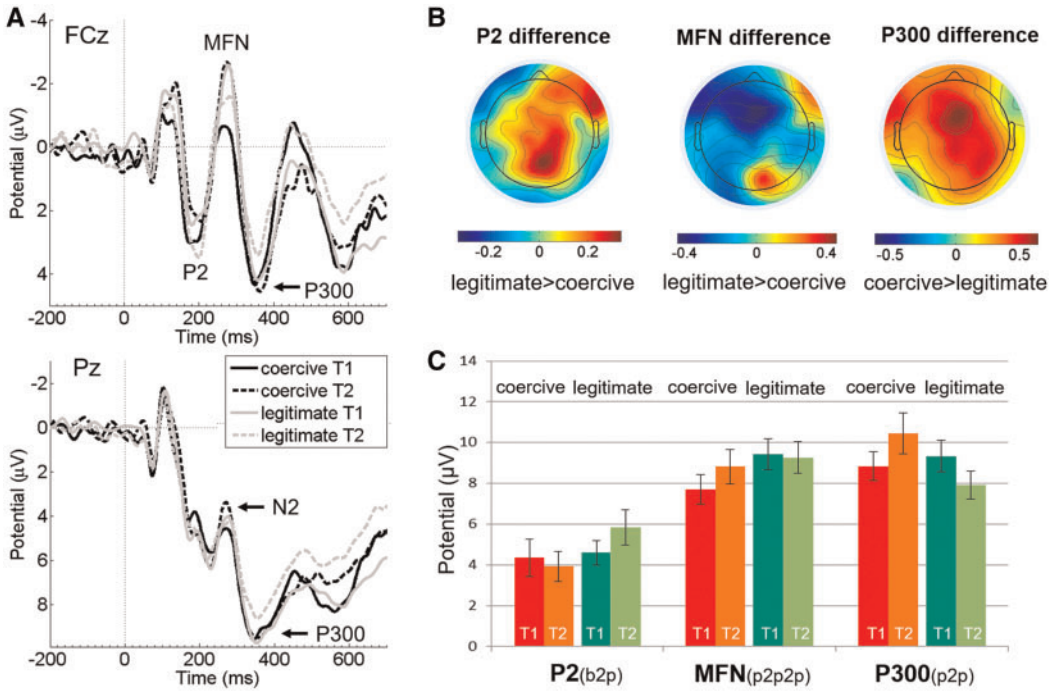

Fig. 2. (A) Time-courses of the P2 and MFN component at FCz (upper panel; also showing P300 peak used for peak quantification [Yeung and Sanfey, 2004]) and the P300 component at Pz (lower panel; also showing N2 peak used for peak quantification [Pfabigan et al., 2011]). Time point zero indicates onset of the flag, tax income, and the 40% tax rate. Negative is drawn upwards per convention. (B) Scalp topographies of the mean difference between coercive and legitimate trials in the P2 time interval (170–210 ms; left side), MFN time interval (260–300 ms; in the middle) and the P300 time interval (300–400 ms, right side). (C) Bar graph depicting mean and SE of P2, MFN and P300 peak values per condition and time point. Please note that absolute values of the three ERPs (in μV) were plotted for a uniform display. Note: b2p = base-to-peak; p2p = peak-to-peak; p2p2p = peak-to-peak-to-peak approaches for ERP quantification.

Behavioral data. Tax payments were not affected by our experimental manipulation (all *P*-values > 0.195). See Supplementary Material 1.2.5 for descriptive statistics.

ERP data. P2 amplitudes were significantly more pronounced under legitimate ($M = 5.20$, $SE = 0.52$) than coercive authority ($M = 4.13$, $SE = 0.58$; $F(1,76) = 5.30$, $P = 0.024$, $\eta_p^2 = 0.07$). As Figure 2a indicates, MFN amplitudes were significantly more pronounced under legitimate ($M = -9.35$, $SE = 0.54$) than coercive authority ($M = -8.27$, $SE = 0.56$; $F(1,76) = 9.43$, $P = 0.003$, $\eta_p^2 = 0.11$). P300 amplitudes were significantly more pronounced under coercive ($M = 9.67$, $SE = 0.62$) than legitimate authority ($M = 8.64$, $SE = 0.52$; $F(1,76) = 6.81$, $P = 0.010$, $\eta_p^2 = 0.08$). See Supplementary Material 1.2.6 for descriptive statistics.

Discussion experiment 2

Experiment 2 confirmed most of the results of experiment 1, apart from the absent differences in tax payments and self-reported rational decision making. However, these absent differences replicate existing findings (Hartl *et al.*, 2015) and might originate from the different settings in both experiments. Concerning ERP data, results show that coercive authority differentially affected attentional processing (P2, P300) and reduced cognitive control demands (MFN) compared to legitimate authority.

General discussion

We investigated whether or not coercive and legitimate centralized institutions elicit comparable cognitive and neuronal processes to reach the final tax payment decision. In experiment 1, coercive authority led to less tax payments, faster decisions and more self-reported rational decision-making, and to less voluntary cooperation than legitimate authority. In experiment 2, no difference in tax payments and self-reported rational decision-making was found. This can be due to differences in experimental procedures, fostering less spontaneous decisions in experiment 2. On the neuronal level, tax authority influenced all ERPs. P2 and MFN amplitudes were more pronounced in legitimate compared to coercive conditions, while P300 amplitudes were more pronounced in coercive than legitimate ones.

In tax experiments which applied a similar paradigm, both coercive and legitimate authority are shown to increase tax payments (Wahl *et al.*, 2010; Hofmann *et al.*, 2014; Hartl *et al.*, 2015). This, however, is based on different motivations. Coercive authority was shown to lead to less trust in institutions, more enforced compliance and less voluntary cooperation than legitimate authority (Hofmann *et al.*, 2014). The different psychological processes underlying both types of authority might explain why some previous studies (Hofmann *et al.*, 2014) as well as experiment 1 show that legitimate authority generates higher tax payments than coercive authority. Enforced compliance elicited by coercive authority reduces tax payments; voluntary cooperation based on legitimate authority increases tax payments (Kastlunger *et al.*, 2012; Gangl *et al.*, 2015). Thus, particularly when individuals decide spontaneously these different motivational processes related to voluntary cooperation and enforced compliance might generate higher tax payments for legitimate than for coercive authority.

In ultimatum game studies, enhanced MFN amplitudes are usually observed after unfair compared to fair offers, and most

likely reflect fairness-related norm enforcement in these situations (Feng *et al.*, 2015). Along these lines, Fehr and Camerer (2007) reported that unfair offers in economic games induce motivational response conflict between economic self-interest and norm enforcement. MFN amplitude variation in the current study might therefore reflect enhanced response conflict induced by legitimate compared to coercive tax authority since evading tax might be perceived as more conflicting in the legitimate than the coercive context because of moral pressure to pursue community goals on costs of egoistic motives. Legitimate authority is likely perceived as positive and as community-serving, which in turn calls for reciprocity (Fehr *et al.*, 1997; Tyler and Fagan, 2008), mirrored by enhanced voluntary cooperation (Hofmann *et al.*, 2014). Thus, MFN and voluntary cooperation data indicate that legitimate authority might make it more difficult to defect than coercive authority. Moreover, we observed significantly larger P2 amplitudes during legitimate than coercive trials. This further corroborates the MFN results. Thus, even before indicating response conflict by MFN variation, legitimate tax authority stimuli induced heightened arousal levels (Carretié *et al.*, 2001) or higher attention capture (Potts, 2004) than coercive tax stimuli.

In contrast to the early evaluation indicated by P2 and MFN components, enhanced P300 amplitudes were observed in coercive compared to legitimate trials. This finding suggests at first glance that coercive authority might yield greater evaluative changes and attention (Cacioppo *et al.*, 1993; Ito *et al.*, 1998; Polich, 2007), which disagrees with a simple calculative cost-benefit heuristic. However, Fabre *et al.* (2015) proposed that larger P300 amplitudes during asset distribution might reflect more automatic processing in line with simple heuristics in working memory (Khader *et al.*, 2011), while decreased amplitudes might reflect a disruption of heuristic-driven processing. For example, P300 amplitudes are larger for fair compared to conflicting unfair and mid-fair proposals (Hewig *et al.*, 2011; Wu *et al.*, 2012; Qu *et al.*, 2013). Also, in dual-task settings, P300 amplitudes are larger in the main task if the second task is an easy compared to a difficult cognitive task (Kramer *et al.*, 1985). Therefore, we belief that legitimate compared to coercive authority is related to a rather difficult decision conflict which disrupts also stimulus categorization (mirrored in P300 decrease).

Faster decisions in experiment 1 in the coercive condition also indicate that a simple calculative cost-benefit analysis might represent an easy-to-apply cognitive heuristic, which is processed faster than the response conflict induced by reciprocity during legitimate authority. Thus, our findings indicate that tax decisions under coercive authority were processed more easily than under legitimate authority.

Self-reported rational decision-making was enhanced under coercive compared to legitimate authority in experiment 1. However, no such difference was observed in experiment 2. Although reaction time and ERP data indicate that the legitimate condition is the more complex one, this self-report finding suggests that the decision complexity under legitimate authority was not consciously perceived by the participants. The different experimental settings might be partly responsible for this. The rather spontaneous tax decision in experiment 1 might have led to more conscious cognitive workload in the coercive than in the legitimate condition. In contrast, the experimentally prolonged tax decision in experiment 2 might have reduced the conscious workload for coercive compared to legitimate authority. Alternatively, the rational decision-making scale might not be able to differentiate between simple calculative cost-benefit analysis and deliberate decisions about the social consequences of one's decision. The present results suggest that both, coercive and legitimate authority, lead to deliberate considerations, but with varying cognitive demands.

Concerning our opposing hypotheses, we found that coercive authority induces simpler calculative cost-benefit analysis instead of more complex social evaluation processes, which in turn make it less conflicting to defect under coercive compared to legitimate authority (Tenbrunsel and Messick, 1999). Surprisingly, the hypothesis that the cost-benefit analysis induced by coercive authority is more complex than the tax decision under legitimate authority is not supported by our data. Relating our findings to previous literature (Coricelli *et al.*, 2010; Dulleck *et al.*, 2016) suggests that coercive authority might induce less deliberate and less emotional demands compared to legitimate authority, which might provoke a straining conflict on reciprocity and fairness accompanied by psychological stress. Thus, coercive and not legitimate authority might provoke more instinctive economic choices (Rubinstein, 2007) and allows a fast calculative response whereas legitimate authority introduces a complex conflict between self- and social interests. Our finding that legitimate authority leads to more fairness-related response conflict might also serve as explanation for why legitimate authority sometimes leads to higher tax payments than coercive authority (e.g. when spontaneous decisions are made such as in experiment 1 and Hofmann *et al.*, 2014). Legitimate more than coercive authority might trigger and explain individuals' tax morale (Frey and Torgler, 2007).

Taxpaying is an excellent example for social decision-making and implies financial decisions which affect both self-interest and other-orientation (Fehr and Camerer, 2007). Several authors suggested that a dual-process system is at work during decision-making reflecting the interplay of self-interest *vs* other-orientation (see Sanfey and Chang, 2008). System 1 is assumed to represent an automatic, fast, effortless, unconscious and slow-learning system implementing automatic and heuristic-based judgements. Contrarily, system 2 is assumed to represent a controlled, slow, effortful and fast-learning system implementing more deliberate reasoning and potentially reappraising input from system 1, and balancing competing interests. Thus, the question arises whether tax decisions under coercive and legitimate authorities are purely based on heuristic-based judgments (system 1) or involve more deliberate reasoning (system 2). The calculative cost-benefit analysis as a heuristic might be used to easily assess self-interests, while social norm enforcement might be reflective of other-orientation. Enhanced MFN amplitudes during legitimate compared to coercive trials might reflect system 1 activity, since MFN variation usually reflects automatic, coarse stimulus evaluation (Hajcak *et al.*, 2006)–indicating enhanced response conflict between self- and other-interests. Subsequent P300 variation indicates that participants might have countered system 1 and tried to costly implement more deliberative reasoning during legitimate trials resulting in diminished P300 amplitudes compared to coercive trials (Fabre *et al.*, 2015).

Limitations of the current study concern the consequences of the necessary compromise between research methods in tax psychology and neuroscience. The experimental setting (within-subject design), number of assessed trials (40) and the assessment of tax decisions (5-point Likert-type scale) deviates from the standard tax paradigms but was necessary to meet the requirements of EEG measurement. However, replication of the effects of coercive and legitimate authority on enforced compliance and voluntary cooperation (Hofmann *et al.*, 2014) show that the

present experiments can be compared with previous ones, providing evidence for the internal validity of the present results.

Future studies on cognitive and neuronal processes involved in tax decisions should examine more realistic settings by analyzing the impact of a combination of high *vs* low coercive with high *vs* low legitimate authority on cooperation (Hartl *et al.*, 2015). Moreover, trust is argued to be as important for cooperation as coercive and legitimate authority (Kirchler *et al.*, 2008).

Cooperation in social systems is organized by centralized institutions such as tax administrations which aim to motivate tax honesty. The present experiments are the first attempt of generating insights into psychological, cognitive and neuronal processes involved in taxpaying under coercive or legitimate authority. Our results also have policy implications: legitimate compared to coercive authority makes tax compliance to a complex social instead of a simple economic problem, thus should be a preferred strategy of tax administrations to foster general tax morale. A better understanding of tax compliance might help to ensure sufficient funds for public goods and might increase public confidence if all taxpayers are perceived to contribute their fair share instead of exploiting the system.

Acknowledgements

We thank Sebastian Fürstenberger, Julia Gäntgen, Valerie Franke and Marianne Gittenberger for their assistance during data collection and analyses. Parts of the data were presented at the EAWOP conference 2013 in Münster, Germany.

Funding

The Austrian Science Fund (FWF; project 24863-G1) supported this study. Participants financial compensation was funded by a scholarship of the Austrian Economic Chamber (WKO) awarded to KG.

Supplementary data

Supplementary data are available at *SCAN* online.

Conflict of interest. All authors declare that this research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Alexander, W.H., Brown, J.W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nature Neuroscience*, **14**(10), 1338–44.

Alexopoulos, J., Pfabigan, D.M., Lamm, C., Bauer, H., Fischmeister, F.P.S. (2012). Do we care about the powerless third? An ERP study of the three-person ultimatum game. *Frontiers in Human Neuroscience*, **6**, 59.

Alexopoulos, J., Pfabigan, D.M., Goschl, F., Bauer, H., Fischmeister, F.P.S. (2013). Agency matters! Social preferences in the three-person ultimatum game. *Frontiers in Human Neuroscience*, **7**, 312.

Alm, J., Jackson, B.R., McKee, M. (1992). Estimating the determinants of taxpayer compliance with experimental data. *National Tax Journal*, **45**, 107–14.

Alm, J., Cherry, T., Jones, M., McKee, M. (2010). Taxpayer information assistance services and tax compliance behavior. *Journal of Economic Psychology*, **31**, 577–86.

Baldassarri, D., Grossman, G. (2011). Centralized sanctioning and legitimate authority promote cooperation in humans. *PNAS*, **108**, 11023–7.

Becker, M.P., Nitsch, A.M., Miltner, W.H., Straube, T. (2014). A single-trial estimation of the feedback-related negativity and its relation to BOLD responses in a time-estimation task. *Journal of Neuroscience*, **34**(8), 3005–12.

Bellebaum, C., Daum, I. (2008). Learning-related changes in reward expectancy are reflected in the feedback-related negativity. *European Journal of Neuroscience*, **27**(7), 1823–35.

Bendor, J., Mookherjee, D. (1990). Norms, third-party sanctions, and cooperation. *Journal of Law, Economics, & Organization*, **6**, 33–63.

Boksem, M.A., De Cremer, D. (2010). Fairness concerns predict medial frontal negativity amplitude in ultimatum bargaining. *Social Neuroscience*, **5**(1), 118–28.

Botvinick, M.M., Carter, C.S., Braver, T.S., Barch, D.M., Cohen, J.D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, **108**(3), 624–52.

Brehm, J.W. (1966). *A Theory of Psychological Reactance*. Oxford, England: Academic Press.

Braithwaite, V. (2003). A new approach to tax compliance. In: Braithwaite, V., editor. *Taxing Democracy*. *Understanding Tax Avoidance and Tax Evasion*, pp. 1–11, Aldershot, UK: Ashgate.

Braithwaite, V. (2003). Dancing with tax authorities: motivational postures and non-compliance actors. In: Braithwaite, V., editor. *Taxing Democracy*, Hants, UK: Ashgate.

Cacioppo, J.T., Crites, S.L., Berntson, G.G., Coles, M. (1993). If attitudes affect how stimuli are processed, should they not affect the event-related brain potential? *Psychological Science*, **4**(2), 108–12.

Camerer, C.F. (2003). Dictator, Ultimatum, and Trust Games, Behavioral Game Theory. Experiments in Strategic Interaction, p 43-116, Princeton University Press/Russell Sage Foundation

Carretié, L., Mercado, F., Tapia, M., Hinojosa, J. (2001). Emotion, attention, and the "negativity bias", studied through eventrelated potentials. *International Journal of Psychophysiology*, **41**, 75–85.

Coricelli, G., Joffily, M., Montmaquette, C., Villeval, M.C. (2010). Cheating, emotions, and rationality: an experiment on tax evasion. *Experimental Economics*, **13**, 226–47.

Dawes, R.M. (1980). Social dilemmas. *Annual Review in Psychology*, **31**, 169–93.

Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., Von Cramon, D.Y., Engel, A.K. (2005). Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *Journal of Neuroscience*, **25**(50), 11730–7.

Delorme, A., Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, **134**(1), 9–21.

Dulleck, U., Fooken, J., Newton, C., Ristl, A., Schaffner, M., Torgler, B. (2016). Tax compliance and psychic costs: behavioral experimental evidence using a physiological marker. *Journal of Public Economics*, **134**, 9–18.

Erard, B., Feinstein, J.S. (1994). The role of moral sentiments and audit perceptions in tax compliance. *Public Finance*, **49**, 70–89.

Fabre, E.F., Causse, M., Pesciarelli, F., Cacciari, C. (2015). Sex and the money–how gender stereotypes modulate economic decision-making: an ERP study. *Neuropsychologia*, **75**, 221–32.

Fehr, E., Camerer, C.F. (2007). Social neuroeconomics: the neural circuitry of social preferences. *Trends in Cognitive Sciences*, **11**(10), 419–27.

Fehr, E., Fischbacher, U. (2002). Why social preferences matter–the impact of non-selfish motives on competition, cooperation and incentives. *Economic Journal*, **112**(478), C1–33.

Fehr, E., Gächter, S., Kirchensteiger, G. (1997). Reciprocity as a contract enforcement device: experimental evidence. *Econometrica*, **65**, 833–60.

Feng, C., Luo, Y.J., Krueger, F. (2015). Neural signatures of fairness-related normative decision making in the ultimatum game: a coordinate-based meta-analysis. *Human Brain Mapping*, **36**(2), 591–602.

Folstein, J.R., Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology*, **45**(1), 152–70.

Frey, B. (1997). A constitution for knaves crowds out civic virtues. *The Economic Journal*, **107**, 1043–53.

Frey, B.S., Torgler, B. (2007). Tax morale and conditional cooperation. *Journal of Comparative Economics*, **35**, 136–59.

Gangl, K., Hofmann, E., Kirchler, E. (2015). Tax authorities' interaction with taxpayers: a conception of compliance in social dilemmas by power and trust. *New Ideas in Psychology*, **37**, 13–23.

Gangl, K., Hofmann, E., de Groot, M., *et al.* (2015). Taxpayers'motivations relating to tax compliance: evidence from two representative samples of Austrian and Dutch self-employed taxpayers. *Journal of Tax Administration*, **1**(2), 15–25.

Gehring, W.J., Willoughby, A.R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, **295**(5563), 2279–82.

Güth, W., Schmittberger, R., Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization*, **3**(4), 367–88.

Güth, W., Levati, M.V., Sutter, M., van der Heijden, E. (2007). Leading by example with and without exclusion power in voluntary contribution experiments. *Journal of Public Economics*, **91**, 1023–42.

Hajcak, G., Moser, J.S., Holroyd, C.B., Simons, R.F. (2006). The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. *Biological Psychology*, **71**(2), 148–54.

Hartl, B., Hofmann, E., Gangl, K., Hartner-Tiefenthaler, M., Kirchler, E. (2015). Does the sole description of a tax authority affect tax evasion?–the impact of descripted coercive and legitimate power. *PLoS One*, **10**, e0123355.

Hechter, M. (2009). Legitimacy in the modern world. *American Behavioral Scientist*, **53**, 279–88.

Hewig, J., Kretschmer, N., Trippe, R.H., *et al.* (2011). Why humans deviate from rational choice. *Psychophysiology*, **48**(4), 507–14.

Hofmann, E., Gangl, K., Kirchler, E., Stark, J. (2014). Enhanced tax compliance through coercive and legitimate power of tax authorities by concurrently diminishing of facilitating trust in authorities. *Law & Policy*, **36**, 290–313.

Hofmann, E., Hartl, B., Gangl, K., Hartner-Tiefenthaler, M., Kirchler, E. (2017). Authorities' coercive and legitimate power: the impact on cognitions underlying cooperation. *Frontiers in Psychology*, **8**, 1–15.

Isakov, A., Rand, D.G. (2012). The evolution of coercive institutional punishment. *Dynamic Games and Applications*, **2**, 97–109.

Ito, T.A., Larsen, J.T., Smith, N.K., Cacioppo, J.T. (1998). Negative information weighs more heavily on the brain: the negativity bias in evaluative categorizations. *Journal of Personality and Social Psychology*, **75**(4), 887–900.

Kastlunger, B., Lozza, E., Kirchler, E., Schabmann, A. (2012). Powerful authorities and trusting citizens: the slippery slope framework and tax compliance in Italy. *Journal of Economic Psychology*, **34**, 36–45.

Khader, P.H., Pachur, T., Meier, S., Bien, S., Jost, K., Rösler, F. (2011). Memory-based decision-making with heuristics involves increased activation of decision re levant memory representations. *Journal of Cognitive Neuroscience*, **23**, 3540–54.

Kirchler, E. (2007). *The Economic Psychology of Tax Behaviour*. Cambridge: Cambridge University Press, UK.

Kirchler, E., Hoelzl, E., Wahl, I. (2008). Enforced versus voluntary tax compliance. The "slippery slope framework". *Journal of Economic Psychology*, **29**, 210–25.

Kirchler, E., Wahl, I. (2010). Tax compliance inventory TAX-I: designing an inventory for surveys of tax compliance. *Journal of Economic Psychology*, **31**, 331–46.

Koslowsky, M., Schwarzwald, J., Ashuri, S. (2001). On the relationship between subordinates' compliance to power sources and organizational attitudes. *Applied Psychology*, **50**, 455–76.

Kramer, A.F., Wickens, C.D., Donchin, E. (1985). Processing of stimulus properties: evidence for dual-task integrity. *Journal of Experimental Psychology: Human Perception and Performance*, **11**, 393–408.

Miltner, W.H.R., Braun, C.H., Coles, M.G.H. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a 'generic' neural system for error detection. *Journal of Cognitive Neuroscience*, **9**(6), 788–98.

Mittone, L. (2006). Dynamic behavior in tax evasion: an experimental approach. *The Journal of Socio-Economics*, **35**, 813–35.

Mossholder, K.W., Bennett, N., Kemery, E.R., Wesolowski, M.A. (2009). Relationships between bases of power and work reactions: the mediational role of procedural justice. *Journal of Management*, **24**, 533–52.

Mulder, L.B., van Dijk, E., De Cremer, D., Wilke, H.A.M. (2006). Undermining trust and cooperation: the paradox of sanctioning systems in social dilemmas. *Journal of Experimental Social Psychology*, **42**, 147–62.

Oldfield, R.C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, **9**(1), 97–113.

Pacini, R., Epstein, S. (1999). The relational and experiential information processing styles to personality, basic beliefs, and the ratio-bias phenomenon. *Journal of Personality and Social Psychology*, **76**(6), 972–87.

Pfabigan, D.M., Alexopoulos, J., Bauer, H., Sailer, U. (2011). Manipulation of feedback expectancy and valence induces negative and positive reward prediction error signals manifest in event-related brain potentials. *Psychophysiology*, **48**(5), 656–64.

Pfabigan, D.M., Sailer, U., Lamm, C. (2015). Size does matter! Perceptual stimulus properties affect event-related potentials during feedback processing. *Psychophysiology*, **52**(9), 1238–47.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology*, **118**(10), 2128–48.

Potts, G.F. (2004). An ERP index of task relevance evaluation of visual stimuli. *Brain and Cognition*, **56**, 5–13.

Qu, C., Wang, Y., Huang, Y. (2013). Social exclusion modulates fairness consideration in the ultimatum game: an ERP study. *Frontiers Human Neuroscience*, **7**, 505.

Raven, B.H., Schwarzwald, J., Koslowsky, M. (1998). Conceptualizing and measuring a power/interaction model of interpersonal influence. *Journal of Applied Social Psychology*, **28**, 307–32.

Rubinstein, A. (2007). Instinctive and cognitive reasoning: a study of response times. *The Economic Journal*, **117**, 1243–59.

Sanfey, A.G., Chang, L.J. (2008). Multiple systems in decision making. *Annals of the New York Academy of Sciences*, **1128**, 53–62.

Shen, L., Dillard, J.P. (2005). Psychometric properties of the Hong psychological reactance scale. *Journal of Personality Assessment*, **85**, 74–81.

Talmi, D., Atkinson, R., El-Deredy, W. (2013). The feedback-related negativity signals salience prediction errors, not reward prediction errors. *Journal of Neuroscience*, **33**(19), 8264–9.

Tenbrunsel, A.E., Messick, D.M. (1999). Sanctioning systems, decision frames, and Cooperation. *Administrative Science Quarterly*, **44**(4), 684–707.

Tyler, T.R. (1997). The psychology of legitimacy: a relational perspective on voluntary deference to authorities. *Personality and Social Psychology Review*, **1**(4), 323–45.

Tyler, T.R., Fagan, J. (2008). Legitimacy and cooperation: why do people help the police fight crime in their communities? *Ohio State Journal of Criminal Law*, **6**, 231–75.

Turner, J.C. (2005). Explaining the nature of power: a three-process theory. *European Journal of Social Psychology*, **35**, 1–22.

Wahl, I., Kastlunger, B., Kirchler, E. (2010). Trust in authorities and power to enforce tax compliance: an empirical analysis of the slippery slope framework. *Law & Policy*, **32**(4), 283–406.

Wu, Y., Hu, J., van Dijk, E., Leliveld, M.C., Zhou, X. (2012). Brain activity in fairness consideration during asset distribution: does the initial ownership play a role? *PLoS ONE*, **7**(6), e39627.

Yeung, N., Botvinick, M.M., Cohen, J.D. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychological Review*, **111**(4), 931–59.

Yeung, N., Sanfey, A.G. (2004). Independent coding of reward magnitude and valence in the human brain. *Journal of Neuroscience*, **24**(28), 6258–64.