**SUPPLEMENTARY MATERIAL**

“Coercive and legitimate authority impact tax honesty: Evidence from behavioral and ERP experiments”

by Katharina Gangl, Daniela M. Pfabigan, Claus Lamm, Erich Kirchler, & Eva Hofmann

1. **SUPPLEMENTARY METHODS**

1.1 Experiment 1

1.1.1 Sample

Participants were recruited at university premises; however, we excluded participants enrolled in psychology to avoid prior knowledge on tax psychology. Sex (χ2(1) = .75, *p* = .82) and age (*t*(77) = .64, *p* = .53) were equally distributed among the two conditions (coercive authority of tax administration followed by legitimate authority [*n* = 39], or legitimate authority followed by coercive authority, [*n* = 41]).

1.1.2 Procedure

The experiment was controlled by E-Prime 2.0 (Psychology Software Tools, Inc., Sharpsburg, PA, USA) using 17” LCD computer screens. Participants were tested in groups of up to 16 individuals.

The fictitious countries differing in coercive and legitimate authority were presented with different names and flags. As names we used Chomland and Sovland which have no meaning in the German language other than ‘land’ signifying ‘country’. As flag colors we used orange with a blue bar and green with a grey bar ensuring no overlapping with real national flags. We used names and national flags to make it more natural and easier for participants to remember and recognize the particularities of each authority. Also, previous studies showed that national flags have an influence on tax compliance (Gangl, Torgler, & Kirchler, in press). Names of the countries and flag colors were counterbalanced across participants.

Authority was manipulated in counterbalanced order. Participants were tested either in a session in which the order was coercive authority followed by legitimate authority or in which the order was legitimate authority followed by coercive authority. Thus, half of the participants experienced coercive authority in their first 40 trials (i.e., tax paying rounds), while the other half experienced legitimate authority in their first 40 trials.

Remuneration of participants was determined by their behavior in the experiment. At the end of the experiment their monetary outcome was determined by one randomly selected taxpaying period in which their earned income in ECU was changed to EUR (exchange rate 1:5.000).

To ensure strong manipulation of authority, a flag and a country name were simultaneously presented with the country descriptions. In addition, participants were reminded of the coercive or legitimate authority every 10th trial. Overall, the experiment lasted about one hour.

To give an example of one of the 40 tax filing trials, one out of five incomes (60.000 ECU, 70.000 ECU, 75.000 ECU, 85.000 ECU, or 90.000 ECU) was presented. Participants were allowed to read information about coercive or legitimate authority, their income, and the tax rate as long as they wanted. Afterwards, they were presented with five possible tax amounts to pay (e.g., with income 60.000 ECU options were 0 ECU, 6.000 ECU, 12.000 ECU, 18.000 ECU, and 24.000 ECU), which corresponded to 0% (fully dishonest), 25%, 50%, 75%, and 100% (fully honest) of the tax due. Participants chose one option via button press. Afterwards, a fixation cross was blended in for a duration of 1000ms until the next trial started.

Figure A1: Example of one tax filing round in experiment 1



Graphical display of the time line of Experiment 1. Each trial started with the presentation of a white fixation cross (1000ms). Subsequently, a short description on tax paying was presented including gross income (60 000 ECU in this example), tax rate and the current country (no time limit). Afterwards, a screen asked how much tax would be paid in the respective tax authority context and the five possible options were presented concurrently (no time limit). Participants indicated their tax decision via button press (buttons 1 – 5) using their right-hand index finger. Afterwards, the fixation cross was presented again for 1000ms and the next trial started.

Statistical analyses: The analyses of variance, correlations and t-tests were conducted with the program IBM SPSS statistics, version 22. The analyses of regressions were conducted with the program R i386 3.3.1 using the packages plm, lmtest, and multiwaycov.

1.1.3 Material

A manipulation check revealed that the coercive tax administration was perceived as significantly more coercive (t(79)=11.83, p<.001, d=1.70; M=6.08, SE=0.14) and less legitimate (t(79)=-10.83, p<.001, d=1.62; M=3.59, SE=0.19) than the legitimate tax administration (coercive authority: M=3.60, SE=0.12; legitimate authority: M=5.36, SE=0.13).

The following results sections demonstrate a complete description of all statistical comparisons. In the main document, we only present significant results.

1.1.4 Self-report data

Self-reported enforced compliance was strongly and significantly affected by the quality of authority, however also the contrast between coercive and legitimate authority induced by the within-subject design, thus the different order of manipulated authority had a small significant effect (interaction: F(1,77)=4.10, p=.046, ηp2=.05; authority: F(1,77)=54.55, p<.001, ηp2=.42, order: F(1,77)=12.34, p=.001, ηp2=.14). The significant interaction effect shows that legitimate authority at time 1 (*M*=4.61, *SE*=0.26) leads to more enforced compliance than legitimate authority at time 2 (*M*=3.11, *SE*=0.24; *t*(78)=-4.20, *p*<.001, *d* = 0.95). The impact of coercive authority on enforced compliance is only by tendency effected by the order (time 1: *M*=5.06, *SE*=0.28; time 2: *M*=5.70, *SE*=0.24; *t*(77)=-1.76, *p*=.083, *d*=0.40). However, comparing the effect sizes of the main effect of authority (ηp2=.42) and the interaction effect (ηp2=.05) shows that the most important impact for enforced compliance is the difference between coercive and legitimate authority.

Voluntary cooperation was higher under legitimate authority than under coercive authority (interaction: (F(1,77)=0.09, p=.764), authority: (F(1,77)=59.51, p<.001, ηp2=.44) order: (F(1,77)=2.85, p=.095). Rational decision making was higher under coercive than under legitimate authority (interaction: F(1,78)=0.59, p=.444, authority: F(1,78)=11.55, p<.001, ηp2=.13), order: (F(1,78)=1.37, p=.245)). Reactance was as expected also higher under coercive authority than under legitimate authority (interaction: (F(1,78)=6.24, p=.015, ηp2=.08), authority: (F(1,78)=81.77, p<.001, ηp2=.51) order: F(1,78)=9.56, p=.003, ηp2=.11)). However, the significant interaction effect indicates (t(78)=-3.66, p<.001, d=0.83) that legitimate authority at time 1 (M=4.41, SE=0.25) leads to more reactance than legitimate authority at time 2 (M=3.16, SE=0.23). The impact of coercive authority on reactance is independent from the order (time 1: M=5.03, SE=0.19; time 2: M= 5.47, SE=0.22; t(78)=-1.52, p=.133). However, comparing the effect sizes of the main effect of authority (ηp2=.51) and the interaction effect (ηp2=.08) shows again that the most important impact for reactance is the difference between coercive and legitimate authority and not the effect induced by the within-subject design.

1.1.5 Behavioral data

Tax payments were independent of the order of authority manipulation (interaction: *F*(1,78)=0.72, *p*=.398; authority: *F*(1,78)=5.96, *p*=.017, ηp2=.07; order: *F*(1,78)=0.61, *p*=.439). Thus, whereas under legitimate authority around 77% of the tax was paid, it was 69% under coercive authority. Means, standard deviations and standard errors by conditions are displayed in Table A1.

Table A1: Experiment 1: Means, standard deviation and standard errors for tax payments

|  |  |  |
| --- | --- | --- |
|  | Coercive authority | Legitimate authority |
|  | T1 n = 39 | T2 n = 41 | T1 n = 41 | T2 n = 39 |
|  | M | SD | SE | M | SD | SE | M | SD | SE | M | SD | SE |
| Tax payments | 3.80 | 1.11 | 0.18 | 3.73 | 1.35 | 0.21 | 3.93 | 1.15 | 0.18 | 4.21 | 0.93 | 0.15 |

To show that the results of the ANOVA concerning tax payments are robust concerning the different dynamic components of our design, we conducted OLS regressions. Table A2 indicates that a regression model in which all single 80 trials of taxpaying are considered lead to the same results as the ANOVA. In model 2, we also control for the different allocated income, the sequence of the 40 trials in each condition and the reminders of the authority. Similar to the ANOVA results, regression results show that legitimate authority leads to higher tax payments than coercive authority.

Table A2: Experiment 1: Robustness check about the experimental effects on tax payments

|  |  |
| --- | --- |
|  | Dependent variable: Tax compliance |
|  | (1) | (2) |
| Coercive vs. legitimate authority | 0.41\*(8.63) | 0.41\*(2.46) |
| Time 1 vs. time 2 | -0.06(-1.23) | -0.06(-0.21) |
| Authority x time | -0.22(-3.28) | -0.22(-0.88) |
| Income  |  | 0.05\*(2.06) |
| Trials (1-40) |  | 0.00(1.25) |
| Reminders |  | 0.07\*(2.19) |
| N | 6261 | 6261 |
| F | 2.68\* | 3.65\*\* |
| R2 | 0.02 | 0.02 |

Note: Standard errors are clustered by individual. Coefficient, t-statistics in parenthesis. Significant levels:\* p < .05, \*\*\* p < .01, \*\*\* p < .001.

Reaction times were significantly affected by the contrast between coercive and legitimate authority (interaction: (F(1,78)=177.63, p<.001, ηp2=.70), authority: (F(1,78)=0.19, p=.661), order: (F(1,78)=2.82, p=.097, ηp2=.04)). Means, standard deviations and standard errors of all conditions are displayed in Table A3.

Table A3: Experiment 1: Means, Standard deviation and standard errors for reaction time

|  |  |  |
| --- | --- | --- |
|  | Coercive authority | Legitimate authority |
|  | T1 n = 39 | T2 n = 41 | T1 n = 41 | T2 n = 39 |
|  | M | SD | SE | M | SD | SE | M | SD | SE | M | SD | SE |
| Reaction time | 3847.89 | 1572.37 | 251.78 | 1784.67 | 887.54 | 138.61 | 3513.01 | 1259.14 | 138.98 | 2230.02 | 867.93 | 196.64 |

To show that the results of the ANOVA concerning reaction time are robust concerning the different dynamic components of our design we conducted OLS regressions. Table A4 indicates that a regression model in which all single 80 trials of taxpaying are considered lead to the same results as the ANOVA. Although the regression model also yields significant main effects of authority and time, the interaction between authority and time is also significant indicating the same results as the ANOVA. In model 2, we also control for the different allocated income, the sequence of the 40 trials in each condition and the reminders of the authority. Similar to the ANOVA results, regression results show that at time 2, coercive authority leads to faster reaction than legitimate authority.

Table A4: Experiment 1: Robustness check about the experimental effects on reaction time

|  |  |
| --- | --- |
|  | Dependent variable: reaction time |
|  | 1 |
|  | (1) | (2) |
| Coercive vs. legitimate authority | -1590.43\*\*\*(-8.05) | -1585.74\*\*\*(-8.03) |
| Time 1 vs. time 2 | -2042.88\*\*\*(-7.362) | -2037.61\*\*\*(-7.356) |
| Authority x time | 3317.72\*\*\*(13.50) | 3313.73\*\*\*(13.50) |
| Income |  | 74.47\*\*\*(3.55) |
| Trials (1-40) |  | -100.16\*\*\*(19.51) |
| Reminders |  | -505.94\*\*\* |
| Number of observations | 6,261 | 6,261 |
| F | 69.08\*\*\* | 71.18\*\*\* |
| R2 | 0.06 | 0.16 |

Note: Standard errors are clustered by individual. Coefficient, t-statistics in parenthesis. Significant levels:\* p < .05, \*\*\* p < .01, \*\*\* p < .001.

1.2 Experiment 2

Please note, the MFN component is often labelled Feedback-Related Negativity component (FRN; Miltner *et al.*, 1997) in non-gambling feedback tasks. The terms FRN and MFN have been used synonymously, we will use MFN throughout the manuscript.

To analyze which of the ERPs is more predictive for tax behavior, we explored Pearson correlations between P2, MFN, P300 amplitudes, and tax behavior.

1.2.1 Sample

Again, we did not include participants enrolled in psychology to avoid prior knowledge on tax psychology. Participants were recruited via social online media. 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 and the University of Vienna. Sex (χ2(1) = .45, *p* = .50) and age (*t*(76) = -0.33, *p* = .74) were equally distributed among the two conditions (coercive authority of tax administration followed by legitimate authority [*n*=38], or legitimate authority followed by coercive authority [*n*=40]).

1.2.2 Procedure

The tax experiment was controlled by E-Prime 2.0 (Psychology Software Tools, Inc., Sharpsburg, PA, USA). Participants were seated in a sound-attenuated and shielded chamber, approximately 70cm in front of a 19” CRT monitor (Sony GDM-F520; 75Hz refresh rate). Overall, EEG data collection lasted about one hour.

Participants performed 10 training trials (without any manipulation of authority) to familiarize them with the experimental set-up and different income levels. After that, the manipulation of coercive authority and legitimate authority was identical to experiment 1.

Compared to experiment 1, however, the difference between incomes levels was reduced (66.000 ECU, 71.000 ECU, 76.000 ECU, 81.000 ECU, or 86.000 ECU). Different income levels were only included to have some variation for participants over the 80 tax paying rounds. Thus, we decided that the different income levels of experiment 1 (min = 60000 ECU; max 90000 ECU) might be too large and be interpreted as either being rich or poor in a current year. To avoid such interpretations which might have caused EEG artefacts in experiment 2, we decided to reduce the income level in experiment 2 to a level we thought more appropriate (min = 66000; max = 86000).

The same self-report questionnaires (asking for enforced compliance, voluntary cooperation, rational decision making and reactance) as in experiment 1 were presented after each of manipulated authority and the following 40 trials. Moreover, again flag colors were counterbalanced across participants.

Similar to experiment 1, the manipulation of authority was conducted in counterbalanced order. However, experiment 1 was a group setting whereas experiment 2 was an individual setting. Thus, only one person was tested at a time. In experiment 2, we first recorded 41 participants starting with coercive authority followed by legitimate authority. Afterwards, we recorded 40 participants starting with legitimate authority followed by coercive authority. Thus, similar to experiment 1, half of the participants experienced coercive authority in their first 40 trials (i.e., tax paying rounds), while the other half experienced legitimate authority in their first 40 trials.

Consequently, half of the participants experienced coercive authority in their first 40 trials (i.e., tax paying rounds), while the other half experienced legitimate authority in their first 40 trials.

We refrained from analyzing reaction times in experiment 2 because we changed the experimental time line slightly from experiment 1 to experiment 2. In particular, we introduced fixed time periods in experiment 2 using slides depicting the current tax income (3000ms) and the tax decision (2000ms) before the participants were actually required to give their response. This is in contrast to experiment 1 which focused on fast decision-making and allowed participants to spend as much time as they wanted reading information regarding the context and their tax income, but allowed them to enter their responses immediately upon being presented with the different tax payment options. These changes were implemented to adapt the tax paradigm of experiment 1 to be used as an ERP experiment.

Figure A2: Example of one tax filing round in experiment 2:



Graphical display of the time line of Experiment 2. Each trial started with the presentation of a white fixation cross (1000ms). Subsequently, gross income and tax rate were presented (3000ms), followed by a screen asking for how much tax would be paid in the respective tax authority context (2000ms). Participants indicated their tax decision via button press (buttons 1 – 5) during the next slide, in which they were presented with five possible options (no time limit). Afterwards, the fixation cross was presented again for a random duration of 1400 – 1600ms.

Because of the changed set-up of experiment 2 allowing assessment of ERP data through fixed times participants had to spent on single slides (Experiment 1 focused on fast decision-making and allowed participants directly to enter their responses upon being presented with the different tax payment options), we refrained from analyzing reaction times in experiment 2.

As in experiment 1, Participants were reminded of the current tax context via presenting short statements describing the respective tax authority. However, in experiment 1 we reminded participants every 10th trials. In experiment 2, we chose to remind our participants every five trials to increase the salience of the current tax authority as much as possible. To avoid boredom in the participants, the reminders rephrased properties of the respective tax authority differently every time. Before starting the experiment, participants were asked several questions regarding the experimental paradigm to ensure task comprehension and compliance.

EEG data were recorded from 57 Ag/AgCl ring electrodes embedded in an equidistant elastic cap (EASYCAP GmbH, Herrsching, Germany; model M10) using a DC amplifier setup (NeuroPrax, neuroConn GmbH, Illmenau, Germany) and sampled at 500 Hz for digital storage. Four additional electrodes were placed above and below the left eye, and on the outer canthi to measure eye-movements. An electrode on the forehead served as online reference. Electrode impedances were kept below 4kΩ using a skin scratching procedure (Picton and Hillyard, 1972).

The EEG experiment lasted one hour. However, including EEG preparation and removal of the EEG cap, participants spent about 2.5 - 3 hours in our laboratory.

Offline EEG data analyses were performed using EEGLAB 11.05.4b (Delorme and Makeig, 2004), implemented in Matlab 7.5.0 (The MathWorks, Inc., Natick, MA, USA). First, a high-pass filter (cutoff frequency 0.05Hz) and a low-pass filter (cutoff frequency 30Hz, roll-off 6dB/octave) were applied. Afterwards, data were re-referenced to linked mastoids before independent component analysis (ICA; Bell and Sejnowski, 1995) was calculated. Individual independent components reflecting vertical or horizontal eye-movements were discarded per participant from the EEG signal. Afterwards, epochs were generated time-locked to the onset of the fictitious income, starting 500ms before income onset and lasting for 3500ms, separately for coercive and legitimate trials. The mean of the first 500 ms served as baseline interval. A semi-automatic artifact removal procedure using EEGLAB algorithms was performed to the epoched data. Trials with voltage values exceeding +/- 75µV (pop\_eegthres) or voltage drifts larger than 50µV (pop\_eegrejtrend) were automatically marked by the software. Trials were rejected in case visual inspection also indicated artifacts. Artifact-free epochs were averaged participant- and condition wise; on average 33.40 trials (SD= 5.87) per condition.

Concerning ERP peak quantification, we chose the following procedure: 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 positive-going P300 component) in the time window 150 – 400ms post income. In detail, we extracted the respective negative-going baseline-to-peak value of the MFN component, the preceding positive-going baseline-to-peak value of the P2 component and the baseline-to-peak value of the following P300 component for each participant and each condition at electrode FCz. Based on Yeung and Sanfey’s (2004) suggestion, we calculated the mean of both positive-going components and subtracted it from the value of the MFN component. Thereby we account for the observation that MFN amplitudes are always influenced by preceding and following positive components. Moreover, via calculating these difference scores, we avoid any spurious results due to individual differences in overall amplitude variation since within our sample, some participants showed rather positive-going amplitude courses while others showed rather negative-going ones.

We applied the same logic to P300 peak extraction. 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 200-600ms post income. In detail, we extracted the respective positive-going baseline-to-peak value of the P300 component and the preceding negative-going baseline-to-peak value of the N2 component. Afterwards, we subtracted N2 values from P300 values (Pfabigan et al., 2011) to account for individual differences and ERP component overlap as with the MFN component.

Both MFN and P300 amplitude values are therefore quantified as difference scores with more pronounced values to be equal to more pronounced amplitude variation (i.e., more negative for MFN component, more positive for P300 component).

The respective time windows for MFN and P300 peak detection are based on available literature and a visual inspection of the current amplitude courses. Using Figure 2 for visual inspection demonstrates that at FCz MFN peaks occurred on average around 300ms post income presentation, P2 peaks roughly around 200ms post income, and P300 peaks almost at 400ms post income. At Pz, the timing looks rather similar, the negative-going N2 peaks are around 300ms post income, the subsequent P300 peaks around 400ms post income. The chosen time windows for peak detection are larger because individual variation in peak latency is rather high. Therefore, we chose larger time windows to capture the respective ERP peaks in all participants. (e.g., Gehring and Willoughby, 2002; Hajcak *et al.*, 2006; Miltner *et al.*, 1997; Pfabigan *et al.,* 2011; Yeungand Sanfey, 2004) and the P300 component (e.g., Bellebaum and Daum, 2008; Pfabigan *et al.*, 2015; Polich, 2007).

To investigate the relationship between taxpaying and ERP components, Pearson correlations were conducted merged for both authorities solely at time 1 (the first 40 trials). We refrained from calculating the correlations for the second 40 trials, since both MFN and P300 components are sensitive to the probability of stimulus occurrence. Change of authorities could be a confound inducing a contrast effect and impacting the subjective perception of stimulus probability.

1.2.3 Material

A manipulation check revealed that the coercive authority was perceived as more coercive (t(77)=11.04, p<.001, d=1.70; M=6.25, SE=0.12) and less legitimate (t(77)=-14.24, p<.001, d=1.62; M=3.75, SE=0.12) than the legitimate authority (coercive authority: M=4.03, SE=0.20; legitimate authority: M=5.74, SE=0.08).

1.2.4 Self-report data

Self-reported enforced compliance was affected by authority (interaction: (*F*(1,76)=6.77, *p*=.011, ηp2=.08), authority: (*F*(1,76)=50.78, *p*<.001, ηp2=.40), order: (*F*(1,76)=5.48, *p*=.022, ηp2=.07)). However, the significant interaction effect, induced by the within-subject design, shows (*t*(76)=-3.27, *p*=.002, *d*=0.75 ) that legitimate authority at time 1 (*M*=4.73, *SE*=0.27) leads to more enforced compliance than legitimate authority at time 2 (*M*=3.49, *SE*=0.27). The impact of coercive authority on enforced compliance is independent from the order (time 1: *M*=5.59, *SE*=0.26; time 2: *M*=5.71, *SE*=0.23*; t*(76)=-0.35, *p*=.73). However, comparing the effect sizes of the main effect of authority (ηp2=.40) and the interaction effect (ηp2=.08) shows that the most important impact for enforced compliance is the difference between coercive and legitimate authority and not the contrast effect induced by the experimental within subject-design.

Voluntary cooperation was significantly higher under legitimate authority than under coercive authority (interaction: (*F*(1,76)=0.13, *p*=.721), authority: (*F*(1,76)=103.11, *p*<.001, ηp2=.58), order: (*F*(1,76)=0.69, *p*=.411)). Rational decision-making was not affected by the manipulation (interaction: (*F*(1,75)=2.68, *p* =.106,), authority: (*F*(1,75)=0.27, *p*=.604), order: (*F*(1,75)=1.04, *p*=.310)). Reactance was affected by authority (interaction: (*F*(1,75)=1.56, *p*=.216), authority: (*F*(1,75)=50.17, *p*<.001, ηp2=.40), order:(*F*(1,75)=0.02, *p*=.894)).

1.2.5 Behavioral data

Tax payments were not significantly different between coercive and legitimate authority, interaction: (*F*(1,76)=1.71, *p*=.195), authority: (*F*(1,76)=0.42, *p*=.521), order: (*F*(1,76)=0.09, *p*=.764); time 1 coercive authority: *M*=3.98, *SE*=0.14, time 1 legitimate authority: *M*=3.95, *SE*=0.16, time 2 coercive authority: *M*=3.98, *SE*=0.18, time 2 legitimate authority: *M*=4.08, *SE*=0.15). Thus, whereas coercive authority elicited about 74% of the tax due it was about 75% for legitimate authority. Means, Standard deviation and standard errors for tax payments are displayed in Table A5.

Table A5: Experiment 2: Means, Standard deviation and standard errors for tax payments

|  |  |  |
| --- | --- | --- |
|  | Coercive authority | Legitimate authority |
|  | T1 n = 38 | T2 n = 40 | T1 n = 40 | T2 n = 38 |
|  | M | SD | SE | M | SD | SE | M | SD | SE | M | SD | SE |
| Tax payments | 3.98 | 0.88 | 0.14 | 3.98 | 1.12 | 0.18 | 3.95 | 1.01 | 0.16 | 4.08 | 0.91 | 0.15 |

To show that the results of the ANOVA concerning tax compliance are robust concerning the different dynamic components of our design we conducted OLS regressions. Table A6 indicates that a regression model in which all single 80 trials of taxpaying are considered lead to the same results as the ANOVA. In model 2, we also control for the different allocated income, the sequence of the 40 trials in each condition and the reminders of the authority. Similar to the ANOVA results, regression results show that coercive and legitimate authority do not differ concerning tax compliance.

Table A6: Experiment 2: Robustness check about the experimental effects on tax payments

|  |  |
| --- | --- |
|  | Dependent variable: tax payments |
|  | (1) | (2) |
| Coercive vs. legitimate authority | 0.18(1.81) | 0.18(1.81) |
| Time 1 vs. time 2 | 0.05(0.24) | 0.05(0.24) |
| Authority x time | -0.21(-1.67) | -0.21(-1.67) |
| Income  |  | -0.02(-0.82) |
| Trials (1-40) |  | -0.00(-0.84) |
| Reminders |  | 0.02(1.00) |
| N | 6320 | 6320 |
| F | 1.20 | 1.77 |
| R2 | 0.00 | 0.00 |

Note: Standard errors are clustered by individual. Coefficient, t-statistics in parenthesis. Significant levels:\* p < .05, \*\*\* p < .01, \*\*\* p < .001

1.2.6 ERP data

Table A7 displays means, standard deviation and standard errors of ERP data.

P2 amplitudes were significantly affected by authority (interaction: *F*(1,76)=0.77, *p*=.383; authority: *F*(1,76)=5.30, *p*=.024, ηp2=.07; order: *F*(1,76)=0.69, *p*=.41). Pearson correlation showed that P2 is not correlated with tax compliance (r=.16, p=.130).

MFN amplitudes were significantly affected by authority (interaction: *F*(1,76)=1.83, *p*=.18), authority: (*F*(1,76)=9.43, *p*=.003, ηp2=.11), order: *F*(1,75)=0.38, *p*=.54). No significant correlation was observed between MFN merged for both authorities and tax payments (*r*=.08, p=.480) or P2 (r=-.11, p=.330).

P300 amplitudes were significantly affected by authority (interaction: *F*(1,76)=0.06, *p*=.814), authority: (F(1,76)=6.81, *p*=.011, ηp2=.08), order: *F*(1,76)=2.05, *p*=.157). Pearson correlation showed a significant relationship between P300 and tax payments merged for both authorities (r=-.23, p=.047). No correlation was found between P300 and P2 (r=.19, p=.100). However, P300 and MFN amplitudes were correlated (r=-.64, p<.001), the correlations between the two ERP components and tax payments failed to differ significantly from each other (Steiger’s *z*-test: z=-1.49, *p*=0.136) – thereby limiting the interpretation of the result regarding P300/tax behavior correlation.

For ERP data it was not applicable to analyze the effect of the different dynamic components of our design with OLS regressions. The trial numbers were too low (and thereby the signal-to-noise ratio too high) for reliable analysis since we presented each income level only eight times per authority. For example, for reliable MFN component analysis, trial numbers are suggested to be higher than 20 (Marco-Pallares et al., 2011).

Table A7: Experiment 2: Means, standard deviation and standard errors of ERP data

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | Coercive authority |  | Legitimate authority |
|  |  |  | T1 (n=38) |  | T2 (n=40) |  | T1 (n=40) |  | T2 (n=38) |
|  |  |  | *M* | *SD* | *SE* |  | *M* | *SD* | *SE* |  | *M* | *SD* | *SE* |  | *M* | *SD* | *SE* |
| FCz | P2 (b2p) |  | 4.35 | 5.64 | 0.92 |  | 3.93 | 4.72 | 0.75 |  | 4.60 | 3.81 | 0.60 |  | 5.84 | 5.25 | 0.85 |
|  | MFN (b2p) |  | -2.60 | 5.95 | 0.97 |  | -3.84 | 6.62 | 1.05 |  | -4.12 | 5.85 | 0.93 |  | -3.75 | 4.54 | 0.74 |
|  | P300 (b2p) |  | 5.85 | 6.08 | 0.99 |  | 6.03 | 5.44 | 0.86 |  | 6.02 | 4.71 | 0.75 |  | 5.18 | 5.53 | 0.90 |
|  | MFN (p2p2p) |  | -7.70 | 4.42 | 0.72 |  | -8.82 | 5.35 | 0.85 |  | -9.43 | 4.82 | 0.76 |  | -9.26 | 4.79 | 0.78 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Pz | N2 (b2p) |  | 3.65 | 4.85 | 0.79 |  | 1.55 | 5.90 | 0.93 |  | 3.00 | 5.19 | 0.82 |  | 3.13 | 3.47 | 0.56 |
|  | P300 (b2p) |  | 12.49 | 6.33 | 1.03 |  | 12.01 | 4.28 | 0.68 |  | 12.34 | 4.59 | 0.73 |  | 11.04 | 5.22 | 0.85 |
|  | P300 (p2p) |  | 8.84 | 4.30 | 0.70 |  | 10.45 | 6.32 | 1.00 |  | 9.33 | 4.84 | 0.77 |  | 7.91 | 4.17 | 0.68 |

Note: b2p: baseline-to-peak; p2p: peak-to-peak; p2p2p: peak-to-peak-to-peak; T1: the first 40 trials of the experiment; T2: the second 40 trials of the experiment

1. **SUPPLEMENTARY DISCUSSION**

**2.1 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 experiment 1 and 2. Experiment 2 was designed to examine ERP data. Thus, compared to experiment 1, individuals were, for example, tested individually (versus in a group), during a longer period of time (approximately 3 hours compared to 1) and importantly, were not allowed to give spontaneous responses. Also we tried to increase the effect of our manipulation and therefore compared to experiment 1, used smaller difference in allocated income and more reminders of the coercive or legitimate authority. It might be that these differences in the experimental setting lead to the different results on tax payments and self-reported rational decision making. Maybe, tax payments are more affected by authority if individuals can give a spontaneous response. However, similar laboratory experiments also showed that coercive authority and legitimate authority to do not differ in their positive impact on tax compliance (Hartl et al., 2015). These results confirm the assumption that both, coercive and legitimate authority foster cooperation, however, based on different underlying psychological motives and processes (Gangl et al., 2015; Kirchler, 2007).

Concerning ERP data, results show that coercive authority leads to differential attentional processing (P2, P300) and reduced cognitive control demands (MFN) compared to legitimate authority. Our MFN results also should be considered in regard to their implications on one of its possible neuronal generators – the anterior cingulate/anterior midcingulate cortex. Regarding current theories on anterior cingulate cortex/anterior midcingulate cortex function, the current experiment is a good example demonstrating the importance of social context factors during performance monitoring. Tax authorities should be considered as social entities which have a considerable impact on each individual. In their literature review, Koban and Pourtois (2014) pointed out that action and performance monitoring is strongly influenced by several affective and social factors – thereby implementing the notion that adaptive behavioural control in our daily life requires us to constantly integrate different information sources. Within their review, the authors propose a neuronal framework to explain the effects of social and affective contextual influences on performance monitoring. They assumed the dorsal medial prefrontal cortex (dMPFC)/anterior midcingulate cortex (aMCC) as the first processing hub dealing with a fast and often automatic detection of errors, of response conflicts, or of reward prediction errors – i.e., reflecting a coarse distinction in favourable vs. unfavourable events. The second processing hub was assumed as the anterior insula, which might be more relevant for more elaborate aspects of performance monitoring. For example, the integration of outcomes, agency, and social context information could take place here, which could lead to more complex emotional experiences such as social emotions. Koban and Pourtois’s hypothesis is based on the idea of integrative processing of affect and performance monitoring outcomes in dorsal medial frontal cortex and the anterior insula as suggested previously (Pessoa, 2008; Shackman et al., 2011). Our current results corroborate the suggestion of the dMPFC/aMCC as the first and rather automatic processing hub regarding social context factors and performance outcomes since MFN amplitude variation was sensitive to the tax authority manipulation. Among others, the aMCC is assumed to be one of the neuronal generators of the MFN/FRN component (Debener *et al.*, 2005). However, the current experimental methods are not suited to go further and investigate subsequent processing steps such as the integration of affect and performance monitoring outcomes reflected in anterior insula activity. Therefore, future studies should aim to integrate research methods with high temporal resolution (such as EEG) and those with high spatial resolution (such as functional magnetic resonance imaging) to better understand the neuro-cognitive processes accompanying tax decisions.

Regarding our P300 results, please note that the current experimental design controlled for known variables affecting P300 amplitudes (stimulus probability and quality, task demands and complexity [Polich, 2007]), thereby limiting their effects on the observed P300 variation in favor of the provided interpretation.

In experiment 2, we further explored whether any ERP component was associated with the behavioral outcome (i.e., averaged tax payments) of the current task. We observed a negative correlation between P300 amplitude variation and tax payments, while correlation coefficients of P2 and MFN components with tax payments were clearly not significant. In order to be able to make strong claims regarding specificity of one correlation compared to a second one, one needs to test whether the correlation coefficients differ from each another. In the current experiment, we were particularly interested whether MFN and P300 components would show distinct correlation patterns with tax behavior. However, as demonstrated by Steiger’s Z test, there was no statistical difference between the two correlation coefficients. Consequently, we refrain from interpreting our correlation results in this regard. Future studies should clarify how robust the finding of a negative association between P300 amplitudes and tax payments is.

**2.2 Discussion – The use of self-report data**

Self-report data allow an introspective measure of the experimental manipulation at hand (such as a quality check of the manipulation) and often demonstrate whether subjective experience of an experimental manipulation is associated with and reflected by physiological correlates or not. In the current experiments, we explicitly asked our participants about their subjective experiences with the two tax authorities. However, as with every assessment tool, self-report has its downsides, which one has to keep in mind. Social desirability might be one critical aspect, lacking introspective abilities another one.

In the current experimental set-up, we assessed self-report measures to indicate whether our experimental manipulation has been implemented successfully or not on a subjective experience level. Indeed, our results corroborate the assumption that the experimental manipulation was successful. We also used self-reported data to interpret the data gathered through reaction time and ERP. Without self-reported data we only would know that there is enhanced activity but we would lack good arguments to interpret the meaning or quality of it (e.g., activity because of enforced compliance or voluntary compliance etc.). We cannot rule out that social desirability or lacking introspective abilities clouded these results. However, the self-report data of a successful experimental manipulation were corroborated by behavioural (Experiment 1) and physiological (Experiment 2) correlates, which are less susceptible to social desirability or other influencing effects. Therefore, we are confident that the use of self-report data added useful information to the current experiments.

References

Bell, A. J. & Sejnowski, T. J. (1995) 'An information-maximization approach to blind

separation and blind deconvolution', *Neural computation*, **7**(6), pp. 1129-1159.

Bellebaum, C. & Daum, I. (2008) 'Learning-related changes in reward expectancy are

 reflected in the feedback-related negativity', *European Journal of Neuroscience*,

 **27**(7), pp. 1823-1835.

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), pp. 11730-11737.

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), pp. 9-21.

Gangl, K., Torgler, B., & Kirchler, E. (in press). Patriotism’s impact on cooperation with the

 state: An experimental study on tax compliance. Political Psychology, doi:

 10.1111/pops.12294

Gehring, W. J. & Willoughby, A. R. (2002) 'The medial frontal cortex and the rapid

 processing of monetary gains and losses', *Science*, **295**(5563), pp. 2279-2282.

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), pp. 148-154.

Koban, L., & Pourtois, G. (2014). Brain systems underlying the affective and social monitoring of actions: An integrative review. *Neurosci Biobehav Rev*, **46P1**, 71-84.

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', *J Cogn Neurosci*, **9**(6), pp. 788-798.

Pessoa, L., 2008. On the relationship between emotion and cognition. *Nat. Rev. Neurosci*.

 **9**, 148–158.

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), pp. 656-

 664.

Pfabigan, D. M., Sailer, U. & Lamm, C. (2015) 'Size does matter! Perceptual stimulus

 properties affect event-related potentials during feedback processing',

 *Psychophysiology*, **52**(9), pp. 1238-1247.

Picton, T. W. & Hillyard, S. A. (1972) 'Cephalic skin potentials in electroencephalography',

 *Electroencephalography and Clinical Neurophysiology*, **33**(4), pp. 419-424.

Polich, J. (2007) 'Updating P300: An integrative theory of P3a and P3b', *Clinical*

 *Neurophysiology*, **118**(10), pp. 2128-2148.

Shackman, A.J., Salomons, T.V., Slagter, H.A., Fox, A.S., Winter, J.J., Davidson, R.J.,

 2011.The integration of negative affect, pain and cognitive control in the

 cingulatecortex. *Nat. Rev. Neurosci*. **12**, 154–167.

Yeung, N. & Sanfey, A. G. (2004) 'Independent coding of reward magnitude and valence in

 the human brain', *Journal of Neuroscience*, **24**(28), pp. 6258-6264.

1. SCENARIOS

|  |  |
| --- | --- |
| Coercive authority scenario | Legitimate authority scenario |
| German original | German original |
| Stellen Sie sich vor, Sie sind Bürger/in im fiktiven Staat Chomland.Die Steuerbehörde in Chomland berechnete, dass das Steueraufkommen etwa 100 Milliarden ECU beträgt. Tatsächlich wurden mehr als 85% der geschätzten Steuern bezahlt.Für dieses gute Ergebnis gab es viele Gründe. Die SteuerprüferInnen greifen hart durch: Kontrollen und Strafen sind der Eckpfeiler der Steuerbehörde. Die Steuerbehörde arbeitet vor allem mit hohen und harten Strafen, um Steuerhinterziehung zu verhindern. Die Steuerbehörde setzt Zwang ein, um SteuerzahlerInnen zu kontrollieren und zu bestrafen. Das Steuerverfahren ist durch abschreckende Maßnahmen gekennzeichnet. Die MitarbeiterInnen der Steuerbehörde strafen hart, damit die SteuerzahlerInnen alle ihre Steuerangelegenheiten offenlegen. Die Steuerbehörde geht davon aus, dass mit Strafen und Kontrollen, die SteuerzahlerInnen zur Steuerehrlichkeit gezwungen werden können: Nur das Androhen negativer Konsequenzen kann Steuerhinterziehung verhindern.Stellen Sie sich die Situation vor, wie Sie als selbstständige/r SteuerzahlerIn in Chomland behandelt werden! Überlegen Sie, wie Sie sich gegenüber der Steuerbehörde fühlen und verhalten! | Stellen Sie sich vor, Sie sind Bürger/in im fiktiven Staat Sovland.Die Steuerbehörde in Sovland berechnete, dass das Steueraufkommen etwa 100 Milliarden ECU beträgt. Tatsächlich wurden mehr als 85% der geschätzten Steuern bezahlt.Für dieses gute Ergebnis gab es viele Gründe. Die SteuerprüferInnen sind Experten: Die Erfahrung und Ausbildung der MitarbeiterInnen sind der Eckpfeiler der Steuerbehörde. Die Steuerbehörde arbeitet vor allem mit effizienten und zielgerichteten Methoden, um Steuerhinterziehung zu verhindern. Die Steuerbehörde setzt Expertenwissen ein, um SteuerzahlerInnen aufzuklären und zu beraten. Das Steuerverfahren ist durch nachvollziehbare Maßnahmen gekennzeichnet. Die MitarbeiterInnen der Steuerbehörde unterstützen professionell, damit die SteuerzahlerInnen alle ihre Steuerangelegenheiten verstehen. Die Steuerbehörde geht davon aus, dass mit Professionalität und Nachvollziehbarkeit, die SteuerzahlerInnen von der Steuerehrlichkeit überzeugt werden können: Nur das Weitergeben von Expertenwissen kann Steuerhinterziehung verhindern.Stellen Sie sich die Situation vor, wie Sie als selbständige/r SteuerzahlerIn in Sovland behandelt werden! Überlegen Sie, wie Sie sich gegenüber der Steuerbehörde fühlen und verhalten! |
| Englisch translation | Englisch translation |
| Imagine, you are a citizen in the fictive state of ChomlandThe tax authority in Chomland calculated the tax revenue with an amount of about 100 billion ECU. In fact, more than 85% of the estimated taxes have been paid.For this good result there were many reasons. The tax auditors take drastic measures: Audits and fines are the cornerstones of the tax authority. The tax authority works primarily with high and harsh punishments to prevent tax evasion. The tax authority uses coercion to control and punish taxpayers. The tax proceeding is characterized by deterrent measures. Employees of the tax authority punish hard, so taxpayers reveal all their tax affairs. The tax authority assumes that taxpayers may be forced to tax compliance by the use of fines and audits: Only the threatening of negative consequences can prevent tax evasion.Imagine the situation, where you will be treated as an independent taxpayer in Chomland! Think about how you feel and behave towards the tax authority! | Imagine, you are a citizen in the fictive state of SovlandThe tax authority in Sovland calculated the tax revenue with an amount of about 100 billion ECU. In fact, more than 85% of the estimated taxes have been paid.For this good result there were many reasons. The tax auditors are experts: The experience and training of employees are the cornerstones of the tax authority. The tax authority works primarily with efficient and targeted methods to prevent tax evasion. The tax authority uses expert knowledge to explain and advise taxpayers. The tax proceeding is characterized by comprehensible measures. Employees of the tax authority do professional support, so that taxpayers understand all their tax affairs. The tax authority assumes that taxpayers can be convinced of tax compliance by professionalism and accountability: Only the sharing of expertise can prevent tax evasion.Imagine the situation, where you will be treated as an independent taxpayer in Sovland! Think about how you feel and behave towards the tax authority! |

1. PSYCHOLOGICAL SCALES

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Experiment 1 | Experiment 2 |

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| Scale | Items | CP | LP | CP | LP |
| Coercive authority | I believe that the tax authority in Sovland/Chomland applies hard punishments on tax evaders. | α = .89 | α =.94 | α = .88 | α = .93 |
|  | I believe that the tax authority in Sovland/Chomland enforces its rules through controls and fines. |  |  |  |  |
|  | I believe that the tax authority in Sovland/Chomland imposes harsh fines if they find an error. |  |  |  |  |
|  | I believe that the tax authority in Sovland/Chomland prosecutes tax evaders with audits and fines |  |  |  |  |
| Legitimate authority | I believe that the tax authority in Sovland/Chomland depends on the cooperation of taxpayers to work efficiently. | α = .85 | α =.88 | α = .83 | α = .73 |
|  | I believe that the tax authority in Sovland/Chomland knows who to counsel taxpayers. |  |  |  |  |
|  | I believe that the tax authority in Sovland/Chomland provides professional counseling of taxpayers. |  |  |  |  |
|  | I believe that the tax authority in Sovland/Chomland provides comprehensible processes in tax collecting. |  |  |  |  |
|  | I believe that the tax authority in Sovland/Chomland makes tax payers feel obliged to cooperate because of the good experiences taxpayers have with the institution. |  |  |  |  |
|  | I believe that the tax authority in Sovland/Chomland cares that the issues of taxpayers are proceeded efficiently and fast. |  |  |  |  |
|  | I believe that the tax authority in Sovland/Chomland has the right to prosecute tax evasion. |  |  |  |  |
|  | I believe that the tax authority in Sovland/Chomland should get a correct tax file if the tax file was submitted incorrectly in the past. |  |  |  |  |
|  | I believe that the tax authority in Sovland/Chomland is appreciated by taxpayers because it uses audits for counseling. |  |  |  |  |
|  | I believe that the tax authority in Sovland/Chomland uses counseling to explain and teach in detail how the tax file should be submitted. |  |  |  |  |
| Enforced compliance | When I pay my taxes as required by the regulations in Sovland/Chomland, I do so because many audits are carried out. | α = .91 | α =.92 | α =.91 | α =.93 |
|  | When I pay my taxes as required by the regulations in Sovland/Chomland, I do so because I know that I will be audited. |  |  |  |  |
|  | When I pay my taxes as required by the regulations in Sovland/Chomland, I do so because the tax office often carries out audits. |  |  |  |  |
| Voluntary cooperation | When I pay my taxes as required by the regulations in Sovland/Chomland, I do so because then the tax office is likely to cooperate with me. | α = .76 | α =.71 | α =.76 | α =.75 |
|  | When I pay my taxes as required by the regulations in Sovland/Chomland, I do so because the tax office is respecting me in the long run as Iong as l admit my mistakes. |  |  |  |  |
|  | When I pay my taxes as required by the regulations in Sovland/Chomland, I do so because the tax office is encouraging to those who have difficulty meeting their obligations through no fault of their own. |  |  |  |  |
| Rational decision making | I solve tax problems through logical thinking. | α =.68 | α =.72 | α =.66 | α =.76 |
|  | I arrive at my tax decisions after deliberate thinking. |  |  |  |  |
|  | I think carful before making an important tax decision. |  |  |  |  |
| Reactance | Regulations of the tax authority in Sovland/Chomland trigger a sense of resistance in me. | α = .73 | α =.83 | α =.86 | α =.83 |
|  | I am frustrated if the tax authorities make me feel depending and not free. |  |  |  |  |
|  | It makes me angry when the tax authority in Sovland/Chomland restricts my freedom of choice. |  |  |  |  |