HAL
open science

# A psychometric investigation of the personality traits underlying individual tax morale

Nicolas Jacquemet, Stephane Luchini, Antoine Malézieux, Jason Shogren

## To cite this version:

Nicolas Jacquemet, Stephane Luchini, Antoine Malézieux, Jason Shogren. A psychometric investigation of the personality traits underlying individual tax morale. 2019. halshs-02008071

**HAL Id: halshs-02008071**

**https://shs.hal.science/halshs-02008071**

Preprint submitted on 5 Feb 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Tax Administration Research Centre

# A psychometric investigation of the personality traits underlying individual tax morale

TARC DISCUSSION PAPER 024 - 19

NICOLAS JACQUEMET, PARIS SCHOOL OF ECONOMICS AND UNIVERSITY PARIS 1 PANTHÉON-SORBONNE

STEPHANE LUCHINI, AIX-MARSEILLE UNIVERSITY

ANTOINE MALEZIEUX, TARC, UNIVERSITY OF EXETER

JASON F SHOGREN, UNIVERSITY OF WYOMING

# A psychometric investigation of the personality traits underlying individual tax morale *

Nicolas Jacquemet† Stéphane Luchini‡ Antoine Malézieux§ Jason F. Shogren¶

January 29, 2019

**Abstract**

Why do people pay taxes? Rational choice theory has fallen short in answering this question. Another explanation, called "tax morale", has been promoted. Tax morale captures the behavioral idea that non-monetary preferences (like norm-submission, moral emotions and moral judgments) might be better determinants of tax compliance than monetary trade-offs. Herein we report on two lab experiments designed to assess whether norm-submission, moral emotions (e.g., affective empathy, cognitive empathy, propensity to feel guilt and shame) or moral judgments (e.g., ethics principles, integrity, and moralization of everyday life) can help explain compliance behavior. Although we find statistically significant correlations of tax compliance behavior with empathy and shame, the economic significance of these correlations are low—more than 80% of the variability in compliance remains unexplained. These results suggest that tax authorities should focus on the institutional context, rather than individual preference characteristics, to handle tax evasion.

**Keywords**: tax evasion, tax morale, morality, personality traits, psychometrics.

**JEL classification**: C9; D03, H26; H31.

*We thank James Alm, Cécile Bazart, Todd Cherry, Taya Cohen, Elisa Darriet, Marine Hainguerlot, Erich Kirchler, Nils Myszkowski, Drazen Prelec, Claire Vandendriessche, Marie-Claire Villeval, and Franck Zenasni for their comments and the helpful discussions from which we benefited in the development of this work. We also thank Kene Boun My for his assistance in running the experiments. The authors thank the *Conseil Régional de Lorraine* for its financial support. We also thank the University of Alaska-Anchorange for their support when writing up this project.

†Paris School of Economics and University Paris 1 Panthéon-Sorbonne. CES, 106 Bd. de l'hôpital, 75013 Paris, France. Nicolas.Jacquemet@univ-paris1.fr

‡Aix-Marseille University (Aix-Marseille School of Economics), CNRS and EHESS. Centre de la Vieille Charitée, 2 Rue de la Charité, 13236 Marseille, France. stephane.luchini@univ-amu.fr

§Tax Administration Research Centre, University of Exeter. Streatham Court, Rennes Drive, Exeter EX4 4ST, United Kingdom. a.e.malezieux@exeter.ac.uk

¶Department of Economics, University of Wyoming, Laramie, WY 82071-3985, United States. JRamses@uwyo.edu

*"We're not interested, frankly, in administering the tax system through fear of penalties."*
Roscoe L. Egger, Jr. (1920–1999), *Wall Street Journal* (Murray, 1984)

# 1 Introduction

Why do people pay taxes given the relatively small risk of an audit and low fines if caught? The answer provided by standard rational choice theory has proven to be inadequate—people who comply must be assumed to have an unrealistically high level of risk aversion (Allingham and Sandmo, 1972). These people would be so risk averse they would never get out of bed in the morning. In response, the "tax morale" literature provides an alternative explanation to the "tax evasion puzzle" (see Torgler, 2002). The tax morale literature specifies this intrinsic motivation to comply by adding non-monetary psychological factors such as individual honesty (Gordon, 1989), a feeling of guilt and shame (Erard and Feinstein, 1994; Andreoni, Erard, and Feinstein, 1998), or a psychic cost of evading (Thomas, 2015) into a person's utility function (see Luttmer and Singhal, 2014, for a survey). Empirical research supports this approach. Declarative tax compliance measures, based on survey data such as the World Value Survey (see e.g. Scholz and Lubell, 1998; Alm and Torgler, 2006; Torgler and Schneider, 2007, 2009; Lago-Peñas and Lago-Peñas, 2010) confirm the importance of attitudinal variables (e.g. levels of trust towards the others, propensity to behave in a generous way, sensibility to equity and equality). Similarly, tax evasion is correlated with physiological measures of emotions, like skin conductance response (Coricelli, Joffily, Montmarquette, and Villeval, 2010) or heart beats rates (Dulleck, Fooken, Newton, Ristl, Schaffner, and Torgler, 2016).

In line with the former IRS Commissioner Roscoe L. Egger Jr. quote, the open question is to pin down those channels allowing to manage the tax system without the "*fear of penalties*". To that end, a promising approach is to widen the scope of the primitives of individual behavior beyond preferences by including the psychological determinants of economic decisions based on personality traits (Borghans, Duckworth, Heckman, and Ter Weel, 2008).[1] Calvet and Alm (2014) initiated the application of this approach to tax evasion by correlating tax compliance behavior and personality measures linked to some moral emotions. The study focuses on empathy and sympathy and only finds few correlations with evasion behavior—a result confirmed by the Principle Component Analysis of Jacquemet, Luchini, Malézieux, and Shogren (2016) based on a larger set of moral emotions. Coricelli, Rusconi, and Villeval (2014) correlate other kinds of moral emotions, shame and guilt, with tax evasion and find that only shame was correlated with the intensity of evasion. A possible interpretation of these results is that previous studies do not account for a broad enough

[1] The relationship between personality traits and behavior in individual tasks, as well the stability of such personality traits in the course of the life-cycle and / or across decision-tasks is the matter of intense debates between the situationist and the behaviorist approach in psychometrics (Almlund, Duckworth, Heckman, and Kautz, 2011; Lefevor, Fowers, Ahn, Lang, and Cohen, 2017). This paper contributes to this debate by relating a narrowly defined, incentivized behavior, to a wide set of personality measures.

set of morality traits.

Herein we step back and extend this literature in three directions: first, we consider together the range of moral emotions that are hypothesized to be involved in tax compliance decisions: empathy, guilt and shame. Second, we also measure norm submission—one's own propensity to be sensitive to external norms of behavior—which might interact with the moral emotions related to norms of tax compliance. Third, moral psychology underlines that while moral emotions define one's sensitiveness to morality, behaving morally also requires moral judgments, i.e. to be able to define what one ought to do (Haidt, 2008). We add measurement of personality traits related to moral judgment.

We rely on lab experiments that allow to observe tax evasion behavior in a controlled environment in which decisions have financial consequences. Since psychometric questionnaires are time consuming and cognitively demanding for participants, we use a planned missing data design (Little and Rhemtulla, 2013) and split the questionnaires between two different experiments. Both experiment elicit compliance towards the same tax rate, in the same decision environment. We add two distinct sets of psychological questionnaires to the decision of income declaration. In Experiment 1, we focus on moral emotions and measure personality traits related to cognitive and affective empathy as well as guilt and shame, as in the existing literature, to which we add a measure of norm submission. In Experiment 2, we turn to moral judgment based on three components: ethics principles, integrity and moral judgments on acts of everyday life. Experiment 2 moreover strengthens the expected relationship between tax morale on tax compliance through a stronger personal identification towards the taxation mechanism (e.g., Wahl, Muehlbacher, and Kirchler, 2010; Lamberton, De Neve, and Norton, 2014). To that end, subjects choose the use of the funds collected: one of the two possible choices replicates Experiment 1, the other offers a similar but different alternative.

Our results suggest that given significant rate of evasion (37% to 49% across experiments), and high heterogeneity in individual scores to personality questionnaires, there exists little association between the compliance and morale. We find most correlations with moral emotions rather than with moral judgments. Affective and cognitive empathy matter to reduce evasion rates; greater guilt and shame sub-scales, however, lead to a greater rate of evasion. Overall, we find at most weak correlations between tax compliance and personality traits related to morality. These results suggest that tax policy should rather focus attention on the institutional environment (such as rules of taxes collection) that will either decrease evasion or make these moral traits salient. The outline of the paper is as follows: Section 2 reviews the foundations of tax morale in light of moral psychology, Section 3 reports the results from Experiment 1 focusing on moral emotions, Section 4 turns to Experiment 2 that studies the correlation of tax compliance with moral judgements, and Section 5 concludes.

## 2 Foundations of tax morale from moral psychology

A growing field within psychology, moral psychology, aims to understand why people behave well and badly (Haidt, 2001; Doris, 2010; Jourdheuil and Petit, 2015). We investigate the role played by participants' morality on their tax compliance behavior in two steps. First, we replicate in a unified framework and generalize the existing literature by measuring morality through moral emotions—what one feels about the morality of his own actions. We then turn to measures of moral judgment—what one ought to do—in the second experiment. This section provides a review of the existing evidence in psychology about the behavioral outcomes related to each personality traits included in our study, from which we derive our main hypotheses.

### 2.1 Morality, moral emotions and social norms

The psychology of moral emotions has emerged from the idea that moral emotions are developed through evolution, to help people choosing the best strategy in human interactions. This gives rise to a strong relationship between emotions and morality, emotions being seen as either serving reason (Frank, 1988), or complementary to it (Damasio, 1994). Prinz and Nichols (2010) distinguish three types of moral emotions: pro-social emotions that promote "morally good behavior" (empathy, sympathy, concern and compassion), self-blamed emotions that evoke negative self-directed feelings (guilt and shame) and other-blamed emotions, i.e. negative feelings directed towards others (contempt, anger, disgust). We choose to include only the first two, pro-social and self-blamed emotions, in our analysis since the third type seems harder to relate to tax evasion.

Regarding pro-social emotions, the empathy-altruism hypothesis (Batson, Dyck, Brandt, Batson, Powell, McMaster, and Griffitt, 1988) predicts that empathetic persons are more altruistic and fair towards others. Calvet and Alm (2014) test the hypothesis of a negative correlation with tax evasion based on psychometric measures of empathy and sympathy. Only sympathy appears positively correlated with tax compliance. The components of self-blamed emotions, shame and guilt, also exhibit contrasted correlations with tax evasion. As regards shame, Coricelli, Joffily, Montmarquette, and Villeval (2010) show an increase in emotional arousal when evaders are informed that their pictures will be shown to other participants. Coricelli, Rusconi, and Villeval (2014) moreover find that the shame proneness scale from the TOSCA-3 test is negatively correlated with the intensity of the fraud after being caught. Experimental evidence on guilt, the other self blamed emotion, is rather mixed. Thurman, John, and Riggs (1984) observe a significant impact of anticipated guilt on tax evasion decisions, but also show that evaders can resort to neutralization strategies to avoid this feeling. This might explain the discrepancies observed in the literature, as Coricelli, Rusconi, and Villeval (2014) for instance fail to find any correlation of tax evasion with the guilt proneness sub-scale from the TOSCA-3.[2] This is confirmed by Dunn,

[2] TOSCA-3 stands for Test of Self-Conscious Affect, third version, published by Tangney, Dearing, Wagner, and Gramzow (2000).

Farrar, and Hausserman (2016), who substantiate an effect of guilt on tax evasion but also show that the effect varies according to guilt cognition.

Social norms have been shown to be an important determinant of tax submission both in the field (Wenzel, 2005; Bobek, Roberts, and Sweeney, 2007) and in the lab (Alm, McClelland, and Schulze, 1999)—see Bobek, Hageman, and Kelliher (2013) for a survey. Norm submission could be a key mediator of the relationship between tax compliance and moral emotions. We include a measure of norm submission so as to elicit the internal pressure to follow the set of implicit rules defined by tax morale.

### 2.2 Morality and moral judgment

A key requirement to recognize a situation as morally problematic is to be able to formulate moral judgment. We collect measures of moral judgment along three dimensions: ethics principles, integrity and moralization of everyday life. These dimensions are in particular part of the Measuring Morality project, a "*nationally-representative survey of adults in the United States aimed at understanding the interrelations among moral constructs, and at exploring moral differences in the U.S. population*".[3]

The tax literature on these themes is too scarce to be conclusive, and focuses on the first: ethics principles. Although tax ethics and tax morale are sometimes seen as overlapping notions (e.g., Wenzel, 2005; McGee, 2011; Maciejovsky, Schwarzenberger, and Kirchler, 2012; Noll, Schnell, and Zdravkovic, 2016), only a few experimental studies investigate tax ethics as a driving force of compliance. Henderson and Kaplan (2005) measure tax ethics using the Multidimensional Ethics Scale, and find a positive correlation with the likelihood of complying in hypothetical scenarios—a result that confirms the one obtained by Reckers, Sanders, and Roark (1994) on participants judging tax evasion as "ethically wrong". Ghosh and Crain (1995) rely on a measure of Machiavellianism to control for tax ethics; they confirm a positive association with compliance. We complement this literature by adding two dimensions to the psychometric measure of moral judgments. The first is integrity, defined as the attachment to ethical principles. It is expected to foster the effect of one's ethics on compliance. The second is a measure of moralization of everyday life that aims to elicit ethics in daily behavior.

## 3 Experiment 1: moral emotions and tax compliance

The first experiment combines a tax evasion game that delivers incentivized measures of tax compliance with psychometric questionnaires of moral emotions.

[3]More information can be found at: www.kenan.ethics.duke.edu/attitudes/resources/measuring-morality/.

### 3.1 Design of the experiment

The standard game used to measure tax compliance behavior in the experimental literature is fairly straightforward: each participant is asked to report income, knowing that declared income will be taxed according to a common knowledge tax rate. The collected tax is deducted from experimental earnings. The target behavior is the share of income that is actually reported. Although the core decision task is standard, many variations in the design can be found in the literature—often associated with uncertain consequences on tax compliance. Our design balances three objectives: we ensure comparability with the existing literature, we generate enough variability in evasion decisions to correlate the outcome with individual covariates, and we enhance the ecological validity of the tax compliance observed in the laboratory.

In Experiment 1, subjects first earn an income through a real effort task.[4] We use a task first introduced by Alm, Cherry, Jones, and McKee (2010), in which the goal is to sort numbers in ascending order from a 3 * 3 matrix filled with digits generated in random order. Earnings are computed based on the time taken to complete the task, as: 150 ECU - (subject's time * 13). The task is repeated 5 times, earned income from this preliminary stage is the sum of earnings from all tasks. Participants then move to the declaration stage. They are asked to "*declare the amount of income they have earned at the previous stage*" (see Cadsby, Maynes, and Trivedi, 2006, on the importance of the way to ask for compliance). They do so using a cursor, which maximal value corresponds to the full income. The tax rate is fixed, common to all participants, and this declaration task is not repeated. In France, the marginal tax rates on 2014 incomes are: 0%, 14%, 30%, 41% and 45%, applied progressively based on the level of income.[5] We use a tax rate equal to 35% that is announced to participants before the beginning of the declaration stage. Declared income determines the taxed, and effectively collected, amounts from each participants' experimental earnings. Collected taxes are used to finance a real life public good: all money is donated to the World Wide Fund for Nature (WWF). To ensure the credibility of the process, donations given to the WWF are officially certified by WWF-certificates that are emailed directly to the participants. It is important to emphasize that there is no audit in this experiment—this allows us to put the spotlight on compliance-based tax morale. Last, the aim to correlate tax evasion and personality traits pleads for the framing of the task.[6] We describe the experiment as a fiscal simulation and the following words are used to describe the progress of the experiment:

[4]The evidence on the effect of windfall money, as compared to earned income, on tax evasion is mixed; see Boylan and Sprinkle (2001); Kirchler, Muehlbacher, Hoelzl, and Webley (2009); Muehlbacher and Kirchler (2009); Boylan (2010); Bühren and Kundt (2013). We favor this choice to strengthen the external validity of our tax evasion measure.

[5]Finance law number 2014-1654, December 29th 2014.

[6]Contextualization of the tax evasion game has be found to have no impact in Alm, McClelland, and Schulze (1992); Swenson (1996); Durham, Manly, and Ritsema (2014) and to undermine tax evasion in Baldry (1986); Wartick, Madeo, and Vines (1999); King and Sheffrin (2002); Mittone (2006); Choo, Fonseca, and Myles (2015). In all cases, evasion rates remain substantial enough to allow an empirical analysis of tax evasion determinants (see, e.g., Wahl, Muehlbacher, and Kirchler, 2010).

Table 1: Glossary from Experiment 1

| Experiment 1 | | |
|---|---|---|
| **Questionnaire** | **Abbrevation** | **Meaning** |
| Concern for Appropriateness Scale (CAS) | _CSV | Cross-Situational Variability of Behavior |
| | _ATSCI | Attention to Social Comparison Information |
| Questionnaire for Cognitive and Affective Empathy (QCAE) | _COG. E | Cognitive Empathy |
| | _PT | Perspective Taking |
| | _OS | Online Simulation |
| | _AFF. E. | Affective Empathy |
| | _EC | Emotion Contagion |
| | _PERIR | Peripheral Responsivity |
| | _PROXR | Proximal Responsivity |
| Guilt And Shame Proneness (GASP) | _Guilt | Guilt |
| | _NBE | Guilt - Negative Behavior Evaluations |
| | _GR | Guilt - Repair |
| | _NSE | Shame - Negative Self Evaluations |
| | _SW | Shame - Withdrawal Responses |

**Note.** Summary of the abbreviations used and their meanings in questionnaires from Experiment 1.

income, income declaration, tax and tax collected.

### 3.1.1 Psychometric measures of moral emotions

At the end of the experiment, subjects answer a socio-demographic questionnaire. They are then presented with three psychometric questionnaires designed to measure individual personality traits related to the propensity to norm-submission, the level of empathy, and the propensity to feel shame and guilt. We compensate the subjects for this last step by adding 5 Euros each to their experimental earnings. For each of these three dimensions, we use a questionnaire validated in the psychometric literature which consists in collecting subjects' reactions toward a set of sentences. Answers are elicited on Likert's scales (ordinally graduated in an ascending order of intensity according to given labels).[7] A glossary of all scales and subscales used in Experiment 1, as well as the corresponding abbreviations we use in the text, can be found in Table 1.

Norm submission is measured thanks to the Concern for Appropriateness Scale (CAS Lennox and Wolfe, 1984). Subjects are asked to express their degree of agreement, according to 6 possible levels of intensity, with 20 statements describing social behaviors.[8] The degree of norm submis-

[7]One of the most used criteria of reliability in this literature is the Cronbach's alpha, that measures the degree of consistence in answers, thanks to the individual variance of answers compared to the total variance (Cronbach, 1951). This measure ranges from 0 to 1, and is increasing with the internal consistency of the questionnaire. An alpha higher than 0.7 is considered as satisfying.

[8]Such as, for example, "*I tend to show different sides of myself to different people*" or "*If I am the least bit*

sion is stronger when the score to this questionnaire is high. This questionnaire is known to be correlated positively with religiosity and risky behavior (Wolfe, Lennox, and Cutler, 1986), as well as behavioral conformism (Johnson, 1984) and the propensity to feel embarrassed (Sabini, Siepmann, Stein, and Meyerowitz, 2000). This questionnaire is also correlated positively with perfectionism, public self-consciousness, social anxiety (Miller, Omens, and Delvadia, 1991) and with harm avoidance (Bachner-Melman, Bacon-Shnoor, Zohar, Elizur, and Ebstein, 2009). It is negatively correlated with self-esteem (Bachner-Melman, Bacon-Shnoor, Zohar, Elizur, and Ebstein, 2009) and emotional stability (Miller, Omens, and Delvadia, 1991).

***Hypothesis 1.*** People scoring high on this scale are norm-submissive and should be more scrupulous when declaring their income.

The feeling of empathy and its two components, affective and cognitive empathy, are measured thanks to the Questionnaire of Cognitive and Affective Empathy (QCAE Reniers, Corcoran, Drake, Shryane, and Völlm, 2011). It features 31 statements with which participants are asked to express their degree of agreement thanks to 4 possible levels.[9] The global score is increasing with the individual level of empathy, and has been shown to be well correlated with the Interpersonal Reactivity Index, another measure of empathy (Michaels, Horan, Ginger, Martinovich, Pinkham, and Smith, 2014). The literature in psychology shows a positive association with pro-social tendencies (Lockwood, Seara-Cardoso, and Viding, 2014) and justice sensitivity (Yoder and Decety, 2014). QCAE scores are also negatively correlated with scales measuring impulsivity, aggression towards others, psychopathy, and Machiavellianism (Reniers, Corcoran, Drake, Shryane, and Völlm, 2011). The correlation with psychopathy is mainly driven by the affective empathy sub-scale (Seara-Cardoso, Dolberg, Neumann, Roiser, and Viding, 2013).

***Hypothesis 2.*** We hypothesize a negative relationship between empathy and tax evasion.

We measure feelings of guilt and shame by the Guilt and Shame Proneness scale (GASP Cohen, Wolf, Panter, and Insko, 2011). These two feelings are distinguished by the context in which they occur—guilt refers to a feeling that arises in a private context, while shame is a reaction to events occurring in a public context. The GASP is made of 16 scenarios in which subjects have to describe the probability to feel one of these two feelings (according to 7 levels graduated from "*Very unlikely*" to "*Very likely*").[10] Cohen, Wolf, Panter, and Insko (2011) show a positive correlation of the guilt scale with psychometric measures of morality and pro social behaviors. People getting high score on this scale are less likely to behave in a non ethical way, to have delinquent behaviors or to engage in counterproductive behaviors towards their companies (Cohen, 2010; Cohen, Panter, and Turan,

*uncertain as to how to act in a social situation, I look to the behavior of others for cues*".

[9]Such as "*I try to look at everybody's side of a disagreement before I make a decision*" or "*I can easily tell if someone else is interested or bored with what I am saying*".

[10]This questionnaire measures the sensitivity to feel these two feelings across scenarios of transgressions, such as "*Your home is messy and unexpected guests knock on your door and invite themselves in. What is the likelihood that you would avoid the guests until they leave?*" or "*You are privately informed that you are the only one in your group that did not make the honor society because you skipped too many days of school. What is the likelihood that this would lead you to become more responsible about attending school?*".

2012; Cohen, Panter, Turan, Morse, and Kim, 2013). Bracht and Regner (2013) show a positive correlation with generosity in a trust game. The shame scale has contrasted results. It is made of two sub-scales: the first, Negative Self-Evaluations, displays the same correlations as those of guilt scale. The second sub-scale (Shame-Withdrawal Responses), by contrast, is positively correlated with non ethical behaviors, delinquency and pro-social attitudes. As a result, we refrain from aggregating the two components of the shame scale, and consider separately the scores obtained at NSE and SW as they refer to two different constructs.[11]

***Hypothesis 3.*** We expect guilt proneness to be negatively correlated with tax evasion along with NSE from shame proneness. SW should be positively correlated with it.

### 3.1.2 Experimental procedure

All experimental sessions took place at the laboratory of the Strasbourg University (LEES) between October 2014 and march 2015.[12] A 5 Euro show-up fee is added to experimental earnings. The empirical analysis relies on three experimental sessions, each having between 19 and 22 participants. Overall, the data is made of 63 participants, including 25 women and 38 men. 59 are students, among them 15 study economics (or a closely related field). We choose to focus on students in this experiment as it is standard to use student pools of subjects for tax evasion experiments. The available papers that compare tax evasion games between students and non-student samples do find less compliance among students but the magnitude of the difference is small (Alm and Malézieux, 2019). In addition, many experimental treatments (for instance, varying audit probability, offering an information service or positive inducement of filings from the tax administration) have been found to alter students' behaviour in the same direction as for non-student samples (Alm, Bloomquist, and McKee, 2015). Last, the questionnaires we borrow from differential psychology are psychometrically validated on student populations. The average participants' age is 23 years old. Each session lasts one hour and the average earnings are 20 Euro (among which 17 Euro are on average earned by participants, and 3 Euro donated to the WWF).

## 3.2 Results

Table 2 reports summary statistics on earned income, compliance behavior and psychometric measures elicited in the experiment.[13] For all outcomes, we observe a high level of inter-individual

[11] The French translation of CAS and QCAE are taken respectively from Myszkowski, Storme, Zenasni, and Lubart (2014) and Myszkowski, Brunet-Gouet, Roux, Robieux, Malézieux, Boujut, and Zenasni (2016). Translation of the GASP has been made by ourselves. For each of the three questionnaires, the sub-scales and their interpretation are presented in Section A in the Appendix.

[12] The recruitment process of the participants makes use of ORSEE (Greiner, 2015). The experiment is computerized using Econplay (www.econplay.fr).

[13] Jacquemet, Luchini, Malézieux, and Shogren (2016) report on a Principal Component Analysis of the psychometric measures delivered by this experiment. The aim of this first analysis is to reduce the variability of the psychometric scores to a fewer number of orthogonal components. This analysis conclude to weak relationship with

variance which allows to test the above hypotheses. The distribution of answers to the CAS is in line with the one observed by Myszkowski, Storme, Zenasni, and Lubart (2014) (on a sample of 634 undergrad students); the results for the QCAE are similar to those of Reniers, Corcoran, Völlm, Mashru, Howard, and Liddle (2012) (on a sample of 24 students) and answers to the GASP are similar to those obtained by Cohen, Wolf, Panter, and Insko (2011) on two different samples (862 representative American adults and 450 undergrad students). The last column of the table provides a measure of the internal consistency of each sub-scale based on observed Cronbach's alphas. They are much higher than 0.70 for the CAS and each of its sub-scales (CSV and ATSCI), a level of consistency that is in line with what has been observed previously in the literature.[14] The consistency of the QCAE measures are a bit lower, and lower than what has been observed in previous implementations.[15] For the GASP, it is common use to study consistency for each sub-scale separately, and to apply a consistency threshold equal to .60 as these sub-scales are scenarios-based and constituted of 4 items each (see Cohen, Wolf, Panter, and Insko, 2011, for a detailed discussion). As compared to other studies using the GASP (Cohen, Wolf, Panter, and Insko, 2011; Howell, Turowski, and Buro, 2012; Schaumberg and Flynn, 2012; Cohen, Panter, Turan, Morse, and Kim, 2013), the NBE sub-scale exceeds most of the time what can be found in the literature (ranging from .67 to .82). Although NSE approaches the threshold of .60, it is just below what can be found in this literature (alphas between .63 and .70). Both GR and SW are well below the threshold, and lower than what can be found in the literature. Such a lack of consistency for some of the GASP sub-scales has also been observed by Howell, Turowski, and Buro (2012); Cohen, Panter, Turan, Morse, and Kim (2013).[16]

### 3.2.1 Compliance behavior and personality traits

As reported at the bottom of Table 2, tax evasion in Experiment 1 is intense with and average declaration rate equal to 49%. It is also widespread, as only one fourth of all participants—16 subjects—declare 100% of their income. Evasion decisions are also much heterogeneous. 5% of the participants (3 subjects) declare zero income, while 25% declare less than 17% of income and 50% less than 42%. It is interesting to note that the correlation between the amount of income earned and the declaration rate is low (equal to -0.023) and non-significant ($p = .855$): there is no evidence of a wealth effect on compliance behavior.

tax compliance. Herein we rather analyze the marginal and joint correlation of each psychometric measure with tax compliance, and compare the results to a broader set of psychometric measures of morality, included in the second experiment.

[14] Cronbach's alpha range from .77 to .90 in Child and Agyeman-Budu (2010); Sabini, Siepmann, Stein, and Meyerowitz (2000); Ragsdale and Brandau-Brown (2005); this last study reports alpha equal to .83 and to .85 for the CSV and the ATSCI sub-scales

[15] Lockwood, Seara-Cardoso, and Viding (2014) finds alpha equal to .87 and .88 for the Cognitive and the Affective Empathy sub-scales; Reniers, Corcoran, Drake, Shryane, and Völlm (2011) reports alpha equal to .85 and .83. for the two COG. E. sub-scales (PT and OS), while they are similar to ours for the AFF. E. sub-scales.

[16] Table 9 in Appendix C provides the 2×2 correlations observed between these measures.

Figure 1: Compliance and psychometric scores in Experiment 1—Univariate analysis

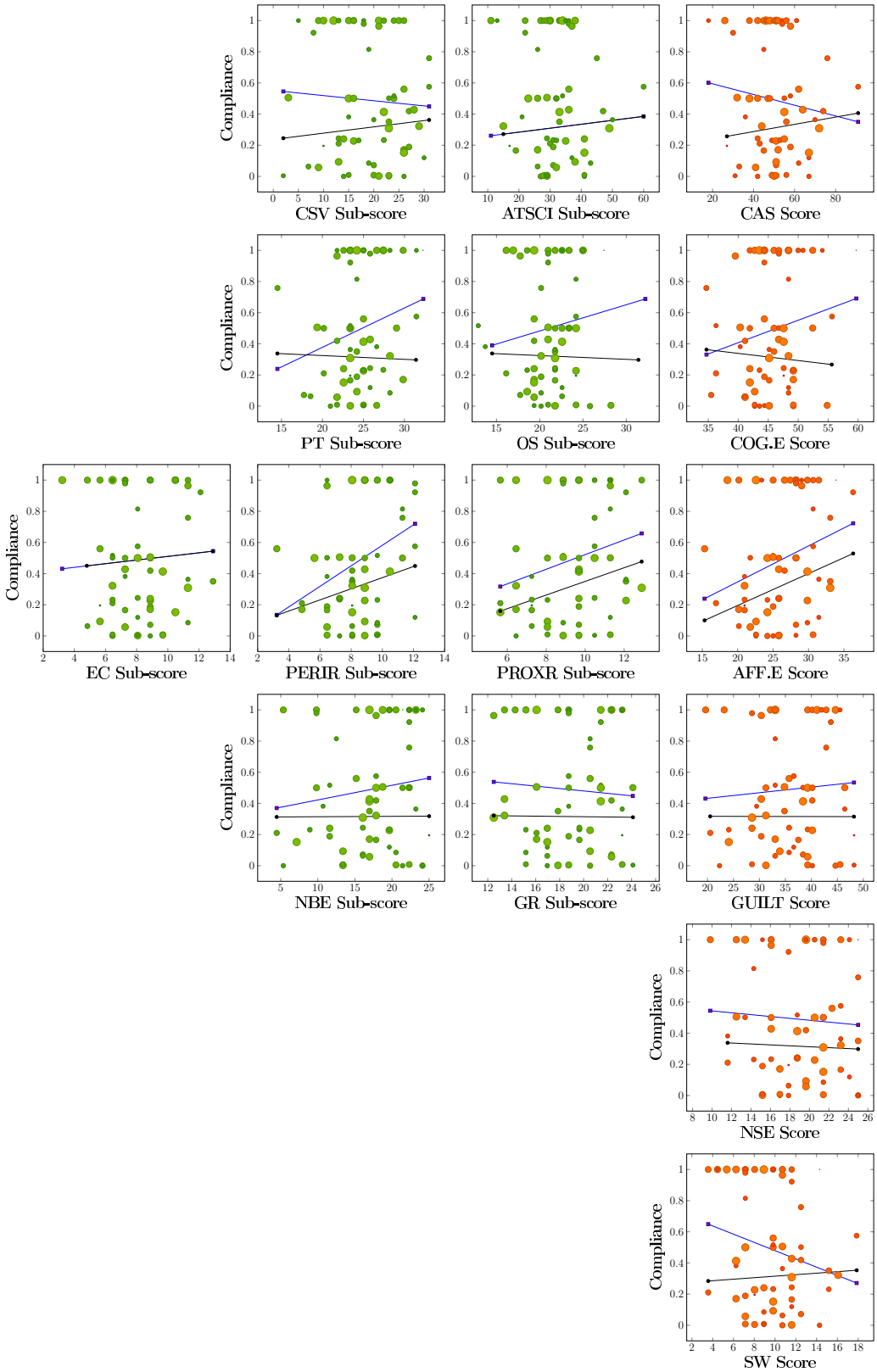


Table 2: Summary statistics on compliance and psychometric measures in Experiment 1

| Variable | Mean | Std. Dev. | Median | Minimum | Maximum | Alpha |
|---|---|---|---|---|---|---|
| CAS | 50.968 | 13.078 | 51 | 18 | 91 | .84 |
| _CSV | 18.841 | 6.996 | 20 | 2 | 31 | .85 |
| _ATSCI | 32.127 | 8.970 | 31 | 11 | 60 | .80 |
| QCAE | 89.063 | 8.811 | 88 | 66 | 116 | – |
| _COG. E. | 56.619 | 5.862 | 57 | 43 | 74 | .74 |
| _PT | 30.270 | 4.009 | 30 | 18 | 40 | .78 |
| _OS | 26.349 | 3.584 | 27 | 16 | 35 | .68 |
| _AFF. E. | 32.444 | 5.477 | 32 | 19 | 45 | .77 |
| _EC | 10.222 | 2.466 | 10 | 4 | 16 | .66 |
| _PROXR | 11.556 | 2.347 | 12 | 7 | 16 | .68 |
| _PERIR | 10.667 | 2.349 | 11 | 4 | 15 | .62 |
| GASP | 72.143 | 11.423 | 74 | 40 | 92 | – |
| _GUILT | 40.222 | 7.701 | 41 | 22 | 54 | – |
| _NBE | 19.222 | 5.709 | 20 | 5 | 28 | .74 |
| _GR | 21.000 | 3.473 | 21 | 14 | 27 | .37 |
| _SHAME | 31.920 | 6.378 | 32 | 17 | 46 | – |
| _NSE | 21.175 | 4.286 | 22 | 11 | 28 | .58 |
| _SW | 10.746 | 3.346 | 11 | 4 | 20 | .37 |
| Income | 356.714 | 87.308 | 362 | 23 | 496 | – |
| Compliance | 0.49 | 0.379 | .41 | 0 | 1 | – |

**Note.** Summary statistics on outcomes from Experiment 1 ($N = 63$ individuals). CAS: Concern for Appropriateness Scale; CSV: Cross-Situational Variability of Behavior; ATSCI: Attention to Social Comparison Information; QCAE: Questionnaire for Cognitive and Affective Empathy; COG. E.: Cognitive Empathy; PT: Perspective Taking; OS: Online Simulation; AFF. E.: Affective Empathy; EC: Emotion Contagion; PERIR: Peripheral Responsivity; PROXR: Proximal Responsivity; GASP: Guilt And Shame Proneness; GUILT: guilt sub-scale from the GASP; NBE: Negative Behavior-Evaluations; GR: Guilt - Repair responses; SHAME: shame sub-scale from the GASP; NSE: Negative Self-Evaluations; SW: Shame - Withdrawal Responses.

Figure 1 provides univariate descriptive evidence on the association between scores to the questionnaires and compliance decisions. Each scale is displayed on the right-hand side along with its sub-scales on the left-hand side. As explained in Section 3.1.1, NSE and SW are considered separately. On each graph, each dot represents one participant of our experiment. The size of the dot and the intensity of its color is proportional to earned income (the bigger and the darker the dot, the higher the income). Two regression lines are drawn: the blue one is based on all observations while the black one focuses only on evaders, i.e. compliance decisions that are not equal to 100% income reporting.

Two main lessons emerge. First, for most measures, the blue and black lines show a weak association between compliance and psychometric scores. Coefficients, confidence intervals, p-values and $R^2$ of these lines can be seen in Table 3. Concerning the whole sample, at the exception

Table 3: Information on the slopes of Figure 1

| | **Full sample** | | | | | | **Sample of compliance < 100%** | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Variable | Coef. | Conf. Inter. | | $p$ | $R^2$ | $r$ | Coef. | Conf. Inter. | | $p$ | $R^2$ | $r$ |
| _CAS | -.003 | -.010 | .004 | .362 | .013 | -.116 | .002 | -.003 | .008 | .443 | .013 | .114 |
| _CSV | -.003 | -.017 | .010 | .635 | .003 | -.060 | .004 | -.007 | .015 | .475 | .011 | .106 |
| _ATSCI | -.005 | -.016 | .005 | .337 | .015 | -.123 | .002 | -.006 | .011 | .576 | .007 | .083 |
| _COG. E. | .011 | -.004 | .027 | .160 | .032 | .179 | -.003 | -.017 | .010 | .605 | .006 | -.077 |
| _PT | .020 | -.003 | .044 | .089 | .046 | .215 | -.001 | .022 | .019 | .854 | .000 | -.027 |
| _OS | .005 | -.021 | .032 | .686 | .002 | .052 | -.007 | -.030 | .015 | .524 | .009 | -.095 |
| _AFF. E. | .018 | .001 | .035 | .033 | .072 | .269 | .016 | .002 | .030 | .018 | .117 | .343 |
| _EC | .009 | -.029 | .048 | .634 | .003 | .061 | .026 | -.007 | .060 | .118 | .053 | .231 |
| _PROXR | .037 | -.002 | .078 | .065 | .054 | .234 | .035 | .002 | .068 | .038 | .092 | .304 |
| _PERIR | .053 | .014 | .092 | .008 | .108 | .329 | .028 | -.002 | .059 | .067 | .072 | .269 |
| _GUILT | .003 | -.009 | .015 | .612 | .004 | .065 | -.000 | -.010 | .010 | .992 | .000 | -.001 |
| _NBE | .008 | -.008 | .025 | .323 | .016 | .126 | .000 | -.014 | .014 | .978 | .000 | .004 |
| _GR | -.006 | -.034 | .020 | .619 | .004 | -.063 | -.000 | -.024 | .022 | .946 | .000 | -.010 |
| _NSE | -.005 | -.027 | .017 | .640 | .003 | -.060 | -.002 | -.022 | .016 | .785 | .001 | -.040 |
| _SW | -.023 | -.052 | .004 | .101 | .043 | -.208 | .004 | -.021 | .029 | .732 | .002 | .051 |

**Note.** Information on the slopes of Figure 1: variables, coefficients, confidence intervals, p-values, $R^2$ and Pearson correlation coefficients $r$. CAS: Concern for Appropriateness Scale; CSV: Cross-Situational Variability of Behavior; ATSCI: Attention to Social Comparison Information; QCAE: Questionnaire for Cognitive and Affective Empathy; COG. E.: Cognitive Empathy; PT: Perspective Taking; OS: Online Simulation; AFF. E.: Affective Empathy; EC: Emotion Contagion; PERIR: Peripheral Responsivity; PROXR: Proximal Responsivity; GASP: Guilt And Shame Proneness; GUILT: guilt sub-scale from the GASP; NBE: Negative Behavior-Evaluations; GR: Guilt - Repair responses; SHAME: shame sub-scale from the GASP; NSE: Negative Self-Evaluations; SW: Shame - Withdrawal Responses.

of PT, AFF. E., PROXR, PERIR and to a lesser extent COG. E. and SW, slopes' coefficients of the blue lines are between -.006 minimum and .009 maximum. None of these variables are significant ($p$ between .323 and .686). The $R^2$ is never higher than .016. Pearson correlation coefficients $r$ (the square root of the $R^2$ in the univariate case) between the scores and compliance are consequently weak (at most .126), indicating that psychometric scores do not explain much of the variance of compliance in this sample. Concerning the sample of participants declaring less than their full incomes, at the exception of AFF. E., PROXR, PERIR and to a lesser extent EC, slopes' coefficient of the blue lines are between -.007 and .004. None of these variables are significant (the $p$-value lies between .443 and .992). The $R^2$ is not higher than 0.013. Pearson correlation coefficients $r$ are again weak (at most .114), confirming that psychometric scores do not explain much of the variance of compliance in both sample. Most of the correlation between scales, sub-scales and mean compliance are extremely low and never significant. This holds when considering the whole sample and only evaders. Only AFF. E., PROXR and PERIR are significant in both settings.

Second, there is a strong discrepancy in the observed univariate association between scores and compliance depending whether the decision to evade is treated separately from its intensity:

Table 4: Experiment 1: Multivariate regressions of compliance decisions on psychometric scores

| | **Extensive margin** | | | | **Intensive margin** | | | |
|---|---|---|---|---|---|---|---|---|
| | Scales | | Sub-scales | | Scales | | Sub-scales | |
| **Variable** | **Coef.** | **(St. E.)** | **Coef.** | **(St. E.)** | **Coef.** | **(St. E.)** | **Coef.** | **(St. E.)** |
| _CAS | -0.018 | (0.020) | – | – | 0.001 | (0.004) | – | – |
| _CSV | – | – | -0.005 | (0.034) | – | – | 0.008 | (0.007) |
| _ATSCI | – | – | -0.022 | (0.031) | – | – | -0.003 | (0.006) |
| _COG. E. | 0.066* | (0.038) | . | . | -0.005 | (0.007) | – | – |
| _PT | – | – | 0.058 | (0.063) | – | – | -0.002 | (0.011) |
| _OS | – | – | 0.022 | (0.068) | – | – | -0.019 | (0.014) |
| _AFF. E. | 0.038 | (0.040) | – | – | 0.019** | (0.008) | – | – |
| _EC | – | – | -0.144 | (0.113) | – | – | 0.003 | (0.023) |
| _PROXR | – | – | 0.151 | (0.149) | – | – | 0.044 | (0.028) |
| _PERIR | – | – | 0.165 | (0.107) | – | – | 0.021 | (0.019) |
| _GUILT | -0.011 | (0.030) | – | – | 0.000 | (0.007) | – | – |
| _NBE | – | – | 0.092 | (0.058) | – | – | 0.001 | (0.010) |
| _GR | – | – | -0.143* | (0.077) | – | – | -0.003 | (0.015) |
| _NSE | 0.010 | (0.063) | -0.016 | (0.069) | -0.011 | (0.012) | -0.010 | (0.013) |
| _SW | -0.134* | (0.074) | -0.198** | (0.088) | -0.001 | (0.014) | -0.004 | (0.016) |
| Intercept | -3.191 | (2.369) | -0.809 | (3.113) | 0.150 | (0.444) | 0.375 | (0.572) |
| (Pseudo) $R^2$ | 0.1796 | | | 0.2952 | | 0.159 | | 0.215 |
| $\chi^2_{(6)}$ | 12.821 | | $\chi^2_{(11)}$ | 21.074 | $F_{(6,40)}$ | 1.263 | $F_{(11,35)}$ | .873 |

**Note.** Left-hand side: Probit model on the decision to evade. The dependent variable equals 1 if declared income is equal to earned income, to 0 otherwise ($N = 63$). Right-hand side: OLS regression on the compliance rate (income declared divided by income earned) estimated on the subsample of evaders ($N = 47$). CAS: Concern for Appropriateness Scale; CSV: Cross-Situational Variability of Behavior; ATSCI: Attention to Social Comparison Information; QCAE: Questionnaire for Cognitive and Affective Empathy; COG. E.: Cognitive Empathy; PT: Perspective Taking; OS: Online Simulation; AFF. E.: Affective Empathy; EC: Emotion Contagion; PERIR: Peripheral Responsivity; PROXR: Proximal Responsivity; GASP: Guilt And Shame Proneness; GUILT: guilt sub-scale from the GASP; NBE: Negative Behavior-Evaluations; GR: Guilt - Repair responses; SHAME: shame sub-scale from the GASP; NSE: Negative Self-Evaluations; SW: Shame - Withdrawal Responses. Significance levels: *10%, **5%, ***1%.

in most cases, the slope of black lines is drastically different from the one of blue lines. As an example, compliance is positively related to PT sub-scale when all subjects are considered (blue line), but is totally flat once full compliers are excluded (black line). This difference points to different determinants of behavior depending on whether the intensive or the extensive margin of compliance behavior is considered.

### 3.2.2 Multivariate analysis

We now turn to parametric models aimed at controlling the correlations between psychometric scores. We distinguish the extensive margin of tax compliance (the decision to evade taxes) from its intensive margin (the intensity of evasion when there is evasion). The first outcome is specified as a 0/1 variable, on which we adjust a Probit model estimated on all individuals. The second

outcome is measured as the ratio between the declared income and the income earned; the effect of psychometric scores is measured through a linear model estimated on evaders.[17] For each of the two dimensions, we estimate two specifications of the models: one based on the general scales—except for NSE and SW that are not aggregated—and one based on the specific sub-scales of each scale. The results are reported in Table 4.

The estimation results, for both the extensive and the intensive margin as well as both the general scales and their sub-scales, confirm the general lesson drawn from univariate analysis: the personality traits related to moral emotions measured by the psychometric questionnaires are weakly associated to the decision to evade taxes. We do observe a few correlations, though, that are moreover different depending on whether the decision to evade or its intensity is considered. Regarding the extensive margin, the Cognitive Empathy scale is significantly (at 10%) and positively linked with the decision to fully comply or evade. This indicates that being more able to figure out and understand the emotional states of others increases the probability to be a full complier. The Shame-Withdrawal sub-scale is also significant, but with a negative sign. Once Guilt is disaggregated into its sub-scales, in column 2, Guilt-Repair appears significant (at the 5% level) with the same sign. This negative correlation is expected for SW as people scoring high on this sub-scale tend to have inappropriate behavior following a transgression (see literature review concerning SW where there is a negative correlation between SW and ethical behaviors). However, the negative correlation is not expected for GR and contradicts the existing literature, as it measures the will to correct or compensate a transgression.[18] Turning to the intensive margin, only the affective empathy scale of the QCAE is significant (at 5% level). Its sign is positive, meaning that those who score higher on this affective empathy scale are declaring more honestly their income.

# 4 Experiment 2: moral judgment and tax compliance

Observed behavior from Experiment 1 shows that *(i)* moral emotions weakly explain the decision to evade taxes and *(ii)* when a correlation does show up, the sign are sometimes highly counter intuitive. In Experiment 2, we assess the robustness of these observations to two variations in the design. First, we turn to an alternative dimension of tax morale by focusing on moral judgment rather than moral emotions. Second, the design of the compliance elicitation task aims to foster the effect of tax morale on tax compliance, by letting participants choose the use of the tax collected. The existing literature on the effect of direct democracy (Dal Bó, Foster, and Putterman, 2010) suggests that when people are involved in setting up the rules that will govern them, they

[17] We specify an OLS model that does not account for the fact that the subsample is selected, because we aim to identify the parameters that are specific to the sub-population that decide to evade taxes.

[18] Considering GR and SW have a low Cronbach's alpha, we are not confident about these conclusions. Dropping them from the analysis does not change anything to the outputs. We can provide the results of these regressions on request.

are subsequently more compliant with these rules. The working hypothesis of Experiment 2 is that letting subjects decide how tax revenues will be spent strengthens the expected relationship between tax morale and tax compliance through a stronger personal identification towards the taxation mechanism.

## 4.1 Design of the experiment

The design of Experiment 2 closely follows Experiment 1—income is first earned through a 9 digit ordering task and taxed at a 35% rate with no penalty on tax evasion. The only exception is the declaration stage. Participants are allowed to choose between two organizations to which the tax collected will be donated: the World Wide Fund for Nature (WWF), the default option used in Experiment 1, or a French organization for the protection of the wildlife, ASPAS (*Association pour la protection des animaux sauvages*). Letting participants choose between only two options is standard in the literature on the direct democracy effect (it is the case e.g. in tax evasion games such as in Alm, Jackson, and McKee, 1993; Alm, McClelland, and Schulze, 1999; Wahl, Muehlbacher, and Kirchler, 2010 but also in public good games, in Messer, Zarghamee, Kaiser, and Schulze, 2007; Messer, Suter, and Yan, 2013; Markussen, Putterman, and Tyran, 2013; Kamei, 2014; Le Sage and Van der Heijden, 2015 or in prisoner's dilemma such as Dal Bó, Foster, and Putterman, 2010 etc.). Authors let participants choose between two options so as to standardize as much as possible their choices, to measure only the impact of letting them choose, and not any other difference such as preferences. To control for the variation in compliance induced by this choice, we need to observe compliance in both states of the world: whether the selected association or the other one actually receives the tax collected. To that end, participants are asked to choose between two possible options: in option 1, the WWF is selected with probability 2/3, while ASPAS will receive the funds with a 1/3 probability; option 2 maintains the same probability distribution but favors ASPAS (selected with 2/3 probability) rather than WWF (1/3). Once an option has been chosen, participants are asked to make two declarations: one if ASPAS is selected, one if is WWF. Participants are then individually informed of the association actually selected to receive their taxes. To ensure the credibility of the donations made in the experiment, the funds given to the WWF and ASPAS are certified thanks to certificates directly issued by the organizations and sent directly to the participants by email.

### 4.1.1 Psychometric measures of moral judgments

As in Experiment 1, subjects are asked to fill in a socio demographic questionnaire at the end of the experiment, followed by psychometric questionnaires (subjects receive a 5 Euro fixed fee as compensation for this step). Three dimensions of moral judgments are included: ethics principles, integrity and the moralization of everyday acts. A glossary of all scales and, subscales used in Experiment 2, as well as the corresponding abbreviations we use in the text, can be found in Table 5.

Table 5: Glossary from Experiment 2

| Experiment 2 | | |
|---|---|---|
| **Questionnaire** | **Abbrevation** | **Meaning** |
| Ethics Position Questionnaire (EPQ) | _Idealism | Idealism |
| | _Relativism | Relativism |
| Integrity Scale (IS) | _Integrity | Integrity |
| Moralization of Everyday Life Scale (MELS) | _F1 - Deception | Factor 1 - Moral weight on the use of deception |
| | _F2 - Norm-violation | Factor 2 - Norm violation |
| | _F3 - Laziness | Factor 3 - Laziness |
| | _F4 - Failure | Factor 4 - Failures to behave the right way |
| | _F5 - Body violations | Factor 5 - Body violations |
| | _F6 - Disgust | Factor 6 - Disgusting behaviors |

**Note.** Summary of the abbreviations used and their meanings in questionnaires from Experiment 2.

The attachment to ethics principles is measured thanks to the Ethics Position Questionnaire (EPQ Forsyth, 1980). The scale is made of two sub-scales, relativism and idealism, both having 10 items.[19] Respondents are asked to report their level of agreement with each statement on a 9 point scale—the higher the score on the relativism sub-scale the higher the rejection of absolute rules; the higher the score on the idealism sub-scale the higher the endorsement of ethical rules. This questionnaire has been extensively used in the last decades, resulting in a large amount of literature (see Davis, Andersen, and Curtis, 2001; Craft, 2013; Meng, Othman, D'Silva, and Omar, 2014, for surveys of its application to business, ethics and academic (dis)honesty). Idealism and relativism are generally correlated with the same outcome behavior, but with reverse signs. For instance, the propensity to use an aggressive business negotiation strategy is related negatively to idealism but positively to relativism (Al-Khatib, Rawwas, Swaidan, and Rexeisen, 2005; Low, Al-Khatib, Vollmers, and Liu, 2007), as is the propensity to morally disengage on a broad range of unethical organizational behaviors (Moore, Detert, Klebe Treviño, Baker, and Mayer, 2012). Idealistic persons tend to see things as being more ethical than less idealistic persons (Singhapakdi, Vitell, and Franke, 1999) and recognize more easily an ethical problem (Dorantes, Hewitt, and Goles, 2006). The reverse relationships are observed for relativistic persons. These personality differences have consequences on behavior. Idealism is negatively correlated with cheating behavior (Sierra and Hyman, 2008), and positively correlated with the likelihood of reporting those who cheated (Smith and Shen, 2013), the rating of academic unethical behaviors as being serious (Etter, Cramer, and Finn, 2006), and stating that reporting peer's cheating is ethical (Barnett, Bass, and

[19] Such as "*People should make certain that their actions never intentionally harm another even to a small degree*" or "*No rule concerning lying can be formulated; whether a lie is permissible or not permissible totally depends upon the situation*".

Brown, 1996). Again, relativism is related to these same behaviors, but with a reverse sign.

***Hypothesis 4.*** We hypothesize a positive correlation between tax compliance and idealism and a negative correlation with relativism.

Integrity, defined as the commitment to ethical principles, is measured by the Integrity Scale (IS Schlenker, 2008). Participant's agreement with 18 statements is elicited on a 5 levels scale.[20] The higher the score, the higher the endorsement of ethics. The score at IS is correlated with a wide range of behaviors and traits. It is positively related with helping others, the frequency of volunteering, the preference for respect, the preference for consistency, role satisfaction and religiosity in Schlenker (2008), with conservatism and life satisfaction in Schlenker, Chambers, and Le (2012), and with religiousness, moral compass (i.e., knowing what is right and wrong), considering lying as unacceptable in Shepperd, Miller, Smith, and Algina (2014). It is also negatively correlated with unethical behavior like plagiarism (Lewis and Zhong, 2011) or cheating (Wowra, 2007).

***Hypothesis 5.*** We expect that a negative association exists between score at IS and tax evasion.

People's assignment of moral weight to common behavior is measured by the Moralization of Everyday Life Scale (MELS Lovett, Jordan, and Wiltermuth, 2012). This scale is made of 30 situations[21] about which participants make moral statements on a 7 levels scale ranging from "*Not wrong at all; has nothing to do with morality*" to "*Very wrong; an extremely immoral action*". The general scale is organized in 6 sub-scales labeled as Factors: Deception (F1, related to the moral weight on the use of deception), Norm violation (F2), Laziness (F3), Failures to behave the right way (F4), Body violations (F5) and Disgust (F6, related to disgusting behaviors). The higher the score, the higher the moralization of everyday life. The existing literature mainly investigates the attitudinal content of the score. The scores to this test significantly explain the variability of the scores to the Crissman (1942) test on everyday moral judgments. Stoeber and Yang (2016) shows that moral perfectionism and general perfectionism explains a part of the variance of the MELS sub-scales. Seeing life thanks to a "dramaturgical perspective" (society is a game where individuals enacts roles) decreases the moralization of everyday behavior (Sullivan, Landau, Young, and Stewart, 2014).[22]

***Hypothesis 6.*** We hypothesize a negative correlation between scores to the MELS test and tax evasion.

[20] Such as "*It is foolish to tell the truth when big profits can be made by lying*" or "*One's principles should not be compromised regardless of the possible gain*".

[21] Such as "*Elizabeth fakes an injury after an automobile accident in order to collect on insurance*" or "*Alexis, a 16-year-old, does not offer her seat on the bus to a disabled old woman*".

[22] The sub-scales and their interpretation are presented in Section B in the Appendix.

Table 6: Summary statistics on compliance and psychometric measures in Experiment 2

| Variable | Mean | Std. Dev. | Median | Minimum | Maximum | Alpha |
|---|---|---|---|---|---|---|
| _Idealism (EPQ) | 66.16 | 12.111 | 67 | 25 | 89 | .83 |
| _Relativism (EPQ) | 53.58 | 12.354 | 53 | 18 | 76 | .74 |
| _Integrity Scale | 60.84 | 8.392 | 60.5 | 41 | 81 | .75 |
| MELS | 113.96 | 26.642 | 112 | 66 | 190 | .90 |
| _F1-Deception | 21.44 | 6.500 | 22 | 8 | 35 | .77 |
| _F2-Norm violation | 27.92 | 6.442 | 29 | 5 | 35 | .81 |
| _F3-Laziness | 8.64 | 4.814 | 7 | 5 | 26 | .78 |
| _F4-Failure | 24.4 | 6.761 | 25 | 5 | 35 | .88 |
| _F5-Body violations | 11.16 | 6.284 | 9 | 5 | 35 | .78 |
| _F6-Disgust | 20.4 | 6.857 | 20.5 | 7 | 34 | .75 |
| Income | 347.70 | 79.810 | 360 | 107 | 483 | – |
| Compliance (for WWF) | 0.366 | 0.311 | .244 | 0 | 1 | – |
| Compliance (for ASPAS) | 0.344 | 0.300 | .222 | 0 | 1 | – |

**Note.** Summary statistics on outcomes from Experiment 2 ($N = 50$). EPQ: Ethics Position Questionnaire; MELS: Moralization of Everyday Life Scale.

### 4.1.2 Experimental procedure

All the experimental sessions took place at the laboratory of the Strasbourg University (LEES) in January 2016.[23] A 5 Euro show-up fee is added to experimental earnings. The empirical analysis relies on two experimental sessions, with 25 participants each. Overall, the data is made of 50 participants, including 25 women and 25 men. All subjects are students, among them 17 study economics (or a closely related field). The average participants' age is 21 years old. Each session lasts one hour and the average earnings are 20 Euro (18 Euros earned on average by participants and 2 Euros donated to one of two organizations).

## 4.2 Results

The top part of Table 6 reports summary statistics on the answers elicited to the EPQ, IS and MELS scales. The observed distributions of psychometric scores are similar to those obtained in seminal studies—Barnett, Bass, and Brown (1996) for the EPQ on a sample of 267 students, Johnson and Schlenker (2007) for the IS on a sample of 1341 participants, Stoeber and Yang (2016) for the MELS on a sample of 243 students. Our participants however scored higher on two sub-scales: F2-Norm violation and F4-Failure, meaning that our participants are judging these domains as being more morale. The last column of the Table provides Cronbach measures of

[23]The recruitment process of the participants uses ORSEE Greiner (2015). The experiment is computerized, thanks to a program from the internet platform Econplay (www.econplay.fr).

internal consistency. For all scales and sub-scales, alpha is higher than a 0.7 threshold. For the EPQ, our alphas are similar to the ones observed in the original Forsyth (1980) study (.80 for idealism and .73 for relativism). For IS, our consistency measure is in the middle of the range observed in the studies reviewed in the previous section.[24] Lastly, for the MELS, the alphas are globally a bit lower than in the original Lovett, Jordan, and Wiltermuth (2012) study, but similar to those reported by the follow-up studies reviewed in the previous section, in which alphas range between .78 and .88.

#### 4.2.1 Compliance behavior and personality traits

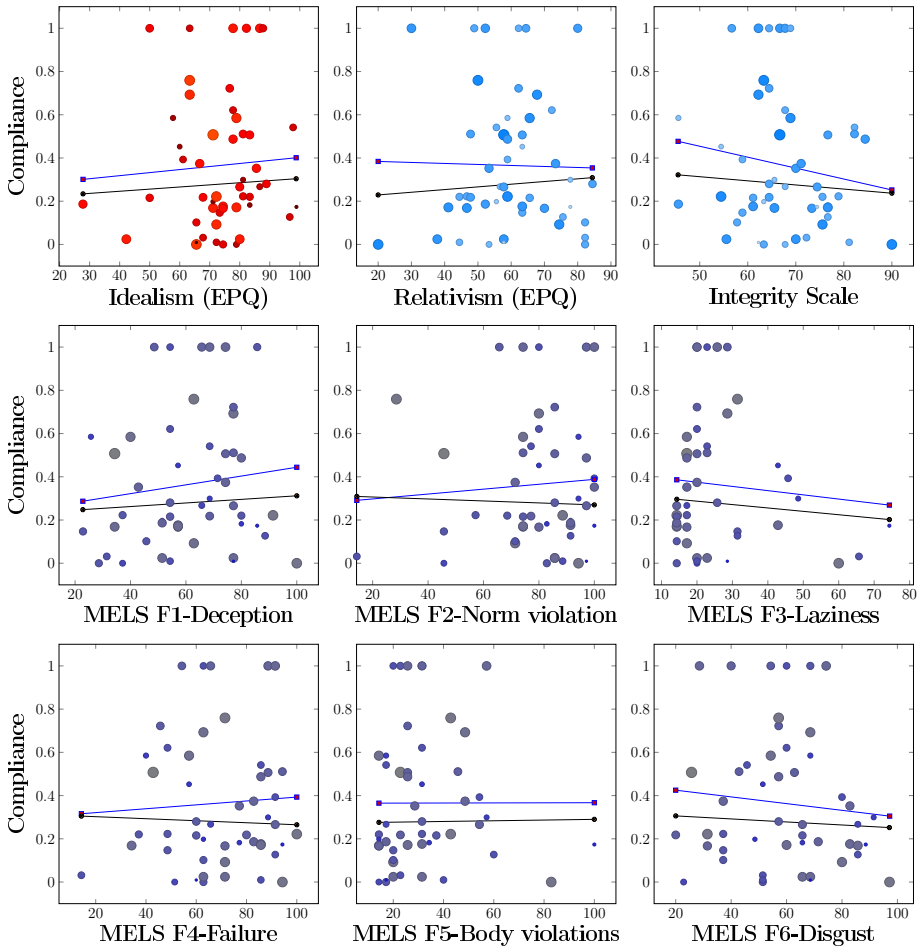
The bottom part of Table 6 describes the distribution of earned income and compliance for both the WWF and ASPAS. The two declarations are highly correlated ($r = .94$) and exhibit similar distributions. To ease the comparison with the results from Experiment 1, we focus on compliance decisions directed towards the WWF.[25] As in Experiment 1, tax evasion is intense (the average declaration rate is 36.58%) and widespread—only one eighth of participants, 6 subjects, declare 100% for their income. Evasion decisions are slightly less heterogeneous than in the first experiment. 6% of the participants (3 subjects) declare zero income, while 25% declare less than 16.70% of their income and 50% less than 24.42%. Last, we again can rule out wealth effects in compliance decisions as the correlation with the level of income is both low (equal to 0.0007), and non-significant ($p = .201$).

In Figure 2, we provide the univariate association between the moral judgments questionnaires and compliance using the same patterns as in Figure 1. Despite a wide inter-individual heterogeneity in both scores and compliance, the regression lines clearly show a lack of association with any of the personality traits included in the experiment. Coefficients, confidence intervals, p-values and $R^2$ of these lines can be seen in Table 7. Concerning the whole sample, slopes' coefficients of the blue lines are between -.005 minimum and .005 maximum. None of the variables are significant ($p$ between .329 and .850). The $R^2$ is never higher than .019, leading to a Pearson correlation coefficients $r$ weak (at most -.140). Concerning the sample of participants declaring less than their full incomes, slopes' coefficients of the black lines are between -.005 and .003. None of the variables are again significant ($p$ between .273 and .973). The $R^2$ is never higher than .027, $r$ is never higher than -.166 and once again, we can conclude that psychometric scores do not explain much of the variance of compliance in this sample. The correlations between scales, sub-scales, and mean compliance are extremely low and never significant. This holds when considering the whole sample and only evaders. All regression lines are flat. In contrast with Experiment 1, we do no find any difference depending on whether we pool all subjects, or focus only on evaders.

[24] The IS alphas are ranging from .84 to maximum .90 across five different samples in the original research from Schlenker (2008), is equal to .83 in Hill, Burrow, Brandenberger, Lapsley, and Quaranto (2010) and to .67 in Shepperd, Miller, Smith, and Algina (2014).

[25] All figures, tables and results are the same when using declaration to the ASPAS rather than WWF. The results are available upon request.

Figure 2: Compliance and psychometric scores in Experiment 2—Univariate analysis




For all sub-scales, the blue regression lines (computed by pooling all subjects) and the black ones (excluding full compliers) are now parallel.

### 4.2.2 Multivariate analysis

Given the similar patterns of univariate associations between the whole sample and the subsample of evaders, we focus our multivariate analysis on pooled regressions on all subjects. The results are provided in Table 8. We estimate two specifications of the model, with varying definitions of the compliance variable. The benchmark model, on the left-hand side, uses the compliance directed towards the WWF. On the right-hand side, we make use of the individual choice of the organization who benefits from collected taxes: we use compliance towards the WWF when the option with higher probability favoring WWF has been chosen and compliance for ASPAS otherwise.

The estimation results confirm the conclusions drawn from univariate analysis: compliance is

Table 7: Information on the slopes of Figure 2

| | Full sample | | | | | | Sample of compliance < 100% | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Variable | Coef. | Conf. Inter. | | $p$ | $R^2$ | $r$ | Coef. | Conf. Inter. | | $p$ | $R^2$ | $r$ |
| _Idealism (EPQ) | .000 | -.006 | .007 | .850 | .000 | .027 | .001 | -.004 | .006 | .613 | .006 | .077 |
| _Relativism (EPQ) | -.001 | -.008 | .005 | .750 | .002 | -.046 | -.000 | -.005 | .005 | .973 | .000 | -.005 |
| _Integrity Scale | -.005 | -.015 | .005 | .329 | .019 | -.140 | -.001 | -.009 | .006 | .651 | .004 | -.069 |
| _F1-Deception | .005 | -.007 | .019 | .388 | .015 | .124 | .003 | -.006 | .013 | .518 | .009 | .098 |
| _F2-Norm violation | .003 | -.009 | .017 | .575 | .006 | .081 | -.001 | -.011 | .008 | .788 | .001 | -.041 |
| _F3-Laziness | -.005 | -.023 | .012 | .545 | .007 | -.087 | -.005 | -.018 | .007 | .417 | .015 | -.123 |
| _F4-Failure | .003 | -.009 | .016 | .594 | .006 | .077 | .001 | -.007 | .011 | .688 | .003 | .061 |
| _F5-Body violations | .000 | -.013 | .014 | .919 | .000 | .014 | .000 | -.010 | .010 | .957 | .000 | .008 |
| _F6-Disgust | -.004 | -.017 | .008 | .493 | .009 | -.099 | -.005 | -.014 | .004 | .273 | .027 | -.166 |

**Note.** Information on the slopes of Figure 2: variables, coefficients, confidence intervals, p-values, $R^2$ and Pearson correlation coefficients $r$. EPQ: Ethics Position Questionnaire; MELS: Moralization of Everyday Life Scale.

weakly related to psychometric measures of personality traits related to moral emotions. When compliance directed towards the WWF is considered, only the score on the integrity scale is significant (at 10%). It shows up with a negative sign, which is highly counterintuitive: more upright people should be less willing to evade. This correlation does not survive the conditioning on the choice of the organization who benefits from collected taxes: in the second model, no sub-scale is significantly correlated with the compliance for the organization selected.

## 5 Conclusion

Tax morale—the socio-psychological determinants of an intrinsic willingness to report income truthfully to the tax authority—is one of the main building blocks on how to better design a tax system (OECD, 2013). In this paper, we combine incentivized measures of tax compliance in the laboratory with psychometric measures of personality traits from moral psychology to investigate their empirical association. While the existing literature has considered varying subsets of individual personality traits, we include a wide set of dimensions of morality, related to norm-submission, moral emotions (cognitive and affective empathy, guilt, shame), and the ability to make moral judgments (ethics principles, integrity and moral judgment of acts of everyday life). We elicit tax compliance in a tax evasion game that favors the influence of tax morale—thanks e.g., to the absence of penalty on evaders, and the use of tax collected to fund a real-world public good—while trying to strengthen the external validity of compliance behavior observed in the laboratory—through e.g., the taxation of a previously earned income.[26]

[26]Laboratory experiments trade the ability to control behavior and collect precise information on a wide set of measures against a lower external validity. See e.g., Torgler (2002) for a discussion of the external validity of laboratory tax evasion games. Since the aim of our experiment is to measure the correlation between tax evasion behavior and personality traits, our results do not rely on the external validity of the quantitative measures of tax

Table 8: Experiment 2: Multivariate regressions of compliance decisions on psychometric scores

| | WWF | | Chosen organization | |
|---|---|---|---|---|
| **Variable** | **Coef.** | **(St. e.)** | **Coef.** | **(St. e.)** |
| _Idealism (EPQ) | 0.007 | (0.005) | 0.006 | (0.005) |
| _Relativism (EPQ) | -0.002 | (0.004) | -0.002 | (0.004) |
| _Integrity Scale | -0.015* | (0.008) | -0.012 | (0.008) |
| _F1-Deception | 0.012 | (0.011) | 0.010 | (0.011) |
| _F2-Norm violation | -0.003 | (0.012) | 0.000 | (0.012) |
| _F3-Laziness | -0.007 | (0.015) | -0.010 | (0.015) |
| _F4-Failure | -0.003 | (0.011) | -0.003 | (0.011) |
| _F5-Body violations | 0.004 | (0.013) | 0.006 | (0.012) |
| _F6-Disgust | -0.005 | (0.010) | -0.005 | (0.009) |
| Intercept | 0.953* | (0.516) | 0.800 | (0.507) |
| $R^2$ | 0.122 | | $R^2$ | 0.109 |
| $F_{(9,40)}$ | .619 | | $F_{(9,40)}$ | .545 |

**Note.** OLS regressions of compliance rate (income declared divided by earned income) on scores to moral judgment questionnaires. Left-hand side: compliance directed towards the WWF. Right-hand side: compliance directed towards the chosen organization. All models are estimated on the whole sample of subjects ($N = 50$). EPQ: Ethics Position Questionnaire; MELS: Moralization of Everyday Life Scale. Significance levels: *10%, **5%, ***1%.

We find that both tax compliance and scores at the psychometric questionnaires exhibit high inter-individual variability. But we observe minimal relationships between the variability of income reporting decisions and the distribution of personality traits, both using univariate analysis and multivariate regression models. A few personality traits turn out significant: the propensity to feel affective and cognitive empathy increases tax compliance. This result underlines the social dimension of behavior related to tax morale: the psychological ability to foresee the effect of one's own actions on the situation and the feelings of others plays an important role in this type of situation. In line with the existing literature, we also find a positive correlation between withdrawing after committing a transgression and tax evasion.

While statistically significant, the economic significance of these correlations is low for our population of subjects—more than 80% of the observed variability of compliance remains unexplained when accounting for either moral emotions or moral judgments. This non-result echoes the evidence collected in the large field experiment by Kleven, Knudsen, Kreiner, Pedersen, and Saez (2011), who also found that personal and socioeconomic characteristics only marginally affect tax compliance for over 40,000 Danish taxpayers. One explanation for such a lack of individual determinism of compliance is that while personality traits help explain the intention to adopt a given behavior, they do not necessarily determine behavior itself (Ajzen, 1985). The open question

evasion, but rather to its covariation with personality traits.

is to understand what features of the income reporting process is better able to link intentions to actions, and to administer the tax system without the fear of penalties. This question is next on our agenda.

## References


Ajzen, I. (1985): "From intentions to actions: A theory of planned behavior," in *Action Control*, pp. 11–39. Berlin Heidelberg: Springer.

Al-Khatib, J., M. Y. Rawwas, Z. Swaidan, and R. J. Rexeisen (2005): "The ethical challenges of global business-to-business negotiations: An empirical investigation of developing countries' marketing managers," *Journal of Marketing Theory and Practice*, 13(4), 46–60.

Allingham, M. G., and A. Sandmo (1972): "Income tax evasion: a theoretical analysis," *Journal of Public Economics*, 1, 323–338.

Alm, J., K. M. Bloomquist, and M. McKee (2015): "On the external validity of laboratory tax compliance experiments," *Economic Inquiry*, 53(2), 1170–1186.

Alm, J., T. Cherry, M. Jones, and M. McKee (2010): "Taxpayer information assistance services and tax compliance behavior," *Journal of Economic Psychology*, 31(4), 577–586.

Alm, J., B. R. Jackson, and M. McKee (1993): "Fiscal exchange, collective decision institutions, and tax compliance," *Journal of Economic Behavior & Organization*, 22(3), 285–303.

Alm, J., and A. Malézieux (2019): "The impact of experimental and individual variables in Tax Evasion Games," Working Paper.

Alm, J., G. H. McClelland, and W. D. Schulze (1992): "Why do people pay taxes?," *Journal of Public Economics*, 48(1), 21–38.

Alm, J., G. H. McClelland, and W. D. Schulze (1999): "Changing the social norm of tax compliance by voting," *Kyklos*, 52(2), 141–171.

Alm, J., and B. Torgler (2006): "Culture differences and tax morale in the United States and in Europe," *Journal of Economic Psychology*, 27(2), 224 – 246.

Almlund, M., A. L. Duckworth, J. Heckman, and T. Kautz (2011): "Personality Psychology and Economics," in *Handbook of the Economics of Education*, vol. 4, chap. Chapter 1, pp. 1–181. Elsevier.

Andreoni, J., B. Erard, and J. Feinstein (1998): "Tax Compliance," *Journal of Economic Literature*, 36(2), 818–860.

Bachner-Melman, R., N. Bacon-Shnoor, A. H. Zohar, Y. Elizur, and R. P. Ebstein (2009): "The psychometric properties of the revised self-monitoring scale (RSMS) and the concern for appropriateness scale (CAS) in Hebrew," *European Journal of Psychological Assessment*, 25(1), 8–15.

Baldry, J. C. (1986): "Tax evasion is not a gamble: A report on two experiments," *Economics Letters*, 22(4), 333–335.

Barnett, T., K. Bass, and G. Brown (1996): "Religiosity, ethical ideology, and intentions to report a peer's wrongdoing," *Journal of Business Ethics*, 15(11), 1161–1174.

Batson, C. D., J. L. Dyck, J. R. Brandt, J. G. Batson, A. L. Powell, M. R. McMaster, and C. Griffitt (1988): "Five studies testing two new egoistic alternatives to the empathy-altruism hypothesis," *Journal of Personality and Social Psychology*, 55(1), 52–77.

Bobek, D. D., A. M. Hageman, and C. F. Kelliher (2013): "Analyzing the role of social norms in tax compliance behavior," *Journal of Business Ethics*, 115(3), 451–468.

Bobek, D. D., R. W. Roberts, and J. T. Sweeney (2007): "The social norms of tax compliance: Evidence from Australia, Singapore, and the United States," *Journal of Business Ethics*, 74(1), 49–64.

Borghans, L., A. L. Duckworth, J. J. Heckman, and B. Ter Weel (2008): "The economics and psychology of personality traits," *Journal of Human Resources*, 43(4), 972–1059.

Boylan, S. J. (2010): "Prior Audits and Taxpayer Compliance: Experimental Evidence on the Effect of Earned Versus Endowed Income," *Journal of the American Taxation Association*, 32(2), 73–88.

Boylan, S. J., and G. B. Sprinkle (2001): "Experimental evidence on the relation between tax rates and compliance: The effect of earned vs. endowed income," *Journal of the American Taxation Association*, 23(1), 75–90.

Bracht, J., and T. Regner (2013): "Moral emotions and partnership," *Journal of Economic Psychology*, 39, 313–326.

Bühren, C., and T. C. Kundt (2013): "Worker or Shirker? Who Evades More Taxes? A Real Effort Experiment," Working Paper.

Cadsby, C. B., E. Maynes, and V. U. Trivedi (2006): "Tax compliance and obedience to authority at home and in the lab: A new experimental approach," *Experimental Economics*, 9(4), 343–359.

Calvet, C. R., and J. Alm (2014): "Empathy, sympathy, and tax compliance," *Journal of Economic Psychology*, 40, 62–82.

Child, J. T., and E. A. Agyeman-Budu (2010): "Blogging privacy management rule development: The impact of self-monitoring skills, concern for appropriateness, and blogging frequency," *Computers in Human Behavior*, 26(5), 957–963.

Choo, C. L., M. A. Fonseca, and G. D. Myles (2015): "Do students behave like real taxpayers in the lab? Evidence from a real effort tax compliance experiment," *Journal of Economic Behavior & Organization*, 124, 102–114.

Cohen, T. R. (2010): "Moral emotions and unethical bargaining: The differential effects of empathy and perspective taking in deterring deceitful negotiation," *Journal of Business Ethics*, 94(4), 569–579.

Cohen, T. R., A. T. Panter, and N. Turan (2012): "Guilt proneness and moral character," *Current Directions in Psychological Science*, 21(5), 355–359.

Cohen, T. R., A. T. Panter, N. Turan, L. Morse, and Y. Kim (2013): "Agreement and similarity in self-other perceptions of moral character," *Journal of Research in Personality*, 47(6), 816–830.

Cohen, T. R., S. T. Wolf, A. T. Panter, and C. A. Insko (2011): "Introducing the GASP Scale: A New Measure of Guilt and Shame Proneness," *Journal of Personality and Social Psychology*, 100(5), 947–966.

Coricelli, G., M. Joffily, C. Montmarquette, and M. C. Villeval (2010): "Cheating, emotions, and rationality: an experiment on tax evasion," *Experimental Economics*, 13(2), 226–247.

Coricelli, G., E. Rusconi, and M. C. Villeval (2014): "Tax evasion and emotions: An empirical test of re-integrative shaming theory," *Journal of Economic Psychology*, 40, 49–61.

Craft, J. L. (2013): "A review of the empirical ethical decision-making literature: 2004–2011," *Journal of Business Ethics*, 117(2), 221–259.

Crissman, P. (1942): "Temporal change and sexual difference in moral judgments," *The Journal of Social Psychology*, 16(1), 29–38.

Cronbach, L. J. (1951): "Coefficient alpha and the internal structure of tests," *Psychometrika*, 16(3), 297–334.

Dal Bó, P., A. Foster, and L. Putterman (2010): "Institutions and behavior: Experimental evidence on the effects of democracy," *American Economic Review*, 100(5), 2205 – 2229.

Damasio, A. R. (1994): *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Putnam.

Davis, M. A., M. G. Andersen, and M. B. Curtis (2001): "Measuring ethical ideology in business ethics: A critical analysis of the ethics position questionnaire," *Journal of Business Ethics*, 32(1), 35–53.

Dorantes, C. A., B. Hewitt, and T. Goles (2006): "Ethical decision-making in an IT context: The roles of personal moral philosophies and moral intensity," in *Proceedings of the 39th Annual Hawaii International Conference*, vol. 8, pp. 206–216.

Doris, J. M. (2010): "Introduction," in *The Moral Psychology Handbook*, ed. by J. M. Doris, and the Moral Psychology Research Group. Oxford: Oxford University Press.

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

Dunn, P., J. Farrar, and C. Hausserman (2016): "The Influence of Guilt Cognitions on Taxpayers' Voluntary Disclosures," *Journal of Business Ethics*, pp. 1–13.

Durham, Y., T. S. Manly, and C. Ritsema (2014): "The effects of income source, context, and income level on tax compliance decisions in a dynamic experiment," *Journal of Economic Psychology*, 40, 220–233.

Erard, B., and J. S. Feinstein (1994): "Honesty and evasion in the tax compliance game," *The RAND Journal of Economics*, 25(1), 1–19.

Etter, S., J. J. Cramer, and S. Finn (2006): "Origins of academic dishonesty: Ethical orientations and personality factors associated with attitudes about cheating with information technology," *Journal of Research on Technology in Education*, 39(2), 133–155.

Forsyth, D. R. (1980): "A taxonomy of ethical ideologies," *Journal of Personality and Social psychology*, 39(1), 175–184.

Frank, R. H. (1988): *Passion within reason*. New York: WW Norton & Company.

Ghosh, D., and T. L. Crain (1995): "Ethical standards, attitudes toward risk, and intentional noncompliance: An experimental investigation," *Journal of Business Ethics*, 14(5), 353–365.

Gordon, J. P. P. (1989): "Individual morality and reputation costs as deterrents to tax evasion," *European Economic Review*, 33(4), 797–805.

Greiner, B. (2015): "Subject pool recruitment procedures: organizing experiments with ORSEE," *Journal of the Economic Science Association*, 1(1), 1–12.

Haidt, J. (2001): "The emotional dog and its rational tail: a social intuitionist approach to moral judgment.," *Psychological review*, 108(4), 814–834.

——— (2008): "Morality," *Perspectives on Psychological Science*, 3(1), 65–72.

Henderson, B. C., and S. E. Kaplan (2005): "An examination of the role of ethics in tax compliance decisions," *Journal of the American Taxation Association*, 27(1), 39–72.

Hill, P. L., A. L. Burrow, J. W. Brandenberger, D. K. Lapsley, and J. C. Quaranto (2010): "Collegiate purpose orientations and well-being in early and middle adulthood," *Journal of Applied Developmental Psychology*, 31(2), 173–179.

Howell, A. J., J. B. Turowski, and K. Buro (2012): "Guilt, empathy, and apology," *Personality and Individual Differences*, 53(7), 917–922.

Jacquemet, N., S. Luchini, A. Malézieux, and J. Shogren (2016): "L'évasion fiscale est-elle un trait de personnalité ? Une évaluation empirique des déterminants psychologiques de la " morale fiscale " [Is tax evasion a personality trait ? An empirical evaluation of psychological determinants of " tax morale "]," *Revue économique*, Forthcoming.

Johnson, M. A. (1984): "Concern for appropriateness scale and behavioral conformity," *Journal of Personality Assessment*, 53(3), 567–574.

Johnson, R., and B. Schlenker (2007): "Assessing the commitment to ethical principles: Psychometric properties of the Integrity Scale," Working Paper.

Jourdheuil, R., and E. Petit (2015): "Émotions morales et comportement prosocial: Une revue de la littérature [Moral emotions and pro-social behavior: A literature review]," *Revue d'économie politique*, 125(4), 499–525.

Kamei, K. (2014): "Democracy and resilient pro-social behavioral change: An experimental study," Working Paper.

King, S., and S. M. Sheffrin (2002): "Tax evasion and equity theory: An investigative approach," *International Tax and Public Finance*, 9(4), 505–521.

Kirchler, E., S. Muehlbacher, E. Hoelzl, and P. Webley (2009): "Effort and aspirations in tax evasion: Experimental evidence," *Applied Psychology*, 58(3), 488–507.

Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez (2011): "Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark," *Econometrica*, 79(3), 651–692.

Lago-Peñas, I., and S. Lago-Peñas (2010): "The determinants of tax morale in comparative perspective: Evidence from European countries," *European Journal of Political Economy*, 26(4), 441–453.

Lamberton, C. P., J.-E. De Neve, and M. I. Norton (2014): "Eliciting taxpayer preferences increases tax compliance," Working Paper.

Le Sage, S., and E. Van der Heijden (2015): "The Effect of Voting on Contributions in a Public Goods Game," Working Paper.

Lefevor, G. T., B. J. Fowers, S. Ahn, S. F. Lang, and L. M. Cohen (2017): "To what degree do situational influences explain spontaneous helping behaviour? A meta-analysis," *European Review of Social Psychology*, pp. 1–30.

Lennox, R. D., and R. N. Wolfe (1984): "Construct validity of the Concern for Appropriateness Scale," *Journal of Personality and Social Psychology*, 46(6), 1349–1364.

Lewis, N. P., and B. Zhong (2011): "The personality of plagiarism," *Journalism & Mass Communication Educator*, 66(4), 325–339.

Little, T. D., and M. Rhemtulla (2013): "Planned missing data designs for developmental researchers," *Child Development Perspectives*, 7(4), 199–204.

Lockwood, P. L., A. Seara-Cardoso, and E. Viding (2014): "Emotion regulation moderates the association between empathy and prosocial behavior," *PloS one*, 9(5).

Lovett, B. J., A. H. Jordan, and S. S. Wiltermuth (2012): "Individual differences in the moralization of everyday life," *Ethics & Behavior*, 22(4), 248–257.

Low, B., J. A. Al-Khatib, S. M. Vollmers, and Y. Liu (2007): "Business-to-business negotiating in China: the role of morality," *Journal of Business & Industrial Marketing*, 22(2), 84–96.

Luttmer, E. F. P., and M. Singhal (2014): "Tax Morale," *Journal of Economic Perspectives*, 28(4), 149–168.

Maciejovsky, B., H. Schwarzenberger, and E. Kirchler (2012): "Rationality versus emotions: The case of tax ethics and compliance," *Journal of Business Ethics*, 109(3), 339–350.

Markussen, T., L. Putterman, and J.-R. Tyran (2013): "Self-organization for collective action: An experimental study of voting on sanction regimes," *The Review of Economic Studies*, 81, 301–324.

McGee, R. W. (2011): *The ethics of tax evasion: Perspectives in theory and practice*. Berlin: Springer Science & Business Media.

Meng, C. L., J. Othman, J. L. D'Silva, and Z. Omar (2014): "Ethical decision making in academic dishonesty with application of modified theory of planned behavior: A review," *International Education Studies*, 7(3), 126–139.

Messer, K. D., J. F. Suter, and J. Yan (2013): "Context Effects in a Negatively Framed Social Dilemma Experiment," *Environmental and Resource Economics*, 55(3), 387–405.

Messer, K. D., H. Zarghamee, H. M. Kaiser, and W. D. Schulze (2007): "New hope for the voluntary contributions mechanism: The effects of context," *Journal of Public Economics*, 91(9), 1783–1799.

Michaels, T. M., W. P. Horan, E. J. Ginger, Z. Martinovich, A. E. Pinkham, and M. J. Smith (2014): "Cognitive empathy contributes to poor social functioning in schizophrenia: evidence from a new self-report measure of cognitive and affective empathy," *Psychiatry Research*, 220(3), 803–810.

Miller, M. L., R. S. Omens, and R. Delvadia (1991): "Dimensions of social competence: Personality and coping style correlates," *Personality and Individual Differences*, 12(9), 955–964.

Mittone, L. (2006): "Dynamic behaviour in tax evasion: An experimental approach," *Journal of Socio-Economics*, 35(5), 813–835.

Moore, C., J. R. Detert, L. Klebe Treviño, V. L. Baker, and D. M. Mayer (2012): "Why employees do bad things: Moral disengagement and unethical organizational behavior," *Personnel Psychology*, 65(1), 1–48.

Muehlbacher, S., and E. Kirchler (2009): "Origin of endowments in public good games: The impact of effort on contributions," *Journal of Neuroscience, Psychology, and Economics*, 2(1), 59–67.

Murray, A. (1984): "Cheating Uncle Sam," *Wall Street Journal*, pp. 1, 12.

Myszkowski, N., E. Brunet-Gouet, P. Roux, L. Robieux, A. Malézieux, E. Boujut, and F. Zenasni (2016): "Is the Questionnaire of Cognitive and Affective Empathy measuring two or five dimensions? Evidence in a French sample," Working Paper.

Myszkowski, N., M. Storme, F. Zenasni, and T. Lubart (2014): "Appraising the duality of self-monitoring: Psychometric qualities of the Revised Self-Monitoring Scale and the Concern for Appropriateness Scale in French.," *Canadian Journal of Behavioural Science*, 46(3), 387–396.

Noll, J., K. Schnell, and S. Zdravkovic (2016): "Tax and Ethics: A Panoramic View," Working Paper.

OECD (2013): "What drives tax morale?," OECD Tax and Development Programme.

Prinz, J. J., and S. Nichols (2010): "Moral emotions," in *The Moral Psychology Handbook*, ed. by J. M. Doris, and the Moral Psychology Research Group. Oxford: Oxford University Press.

Ragsdale, J. D., and F. E. Brandau-Brown (2005): "Individual differences in the use of relational maintenance strategies in marriage," *The Journal of Family Communication*, 5(1), 61–75.

Reckers, P. M., D. L. Sanders, and S. J. Roark (1994): "The influence of ethical attitudes on taxpayer compliance," *National Tax Journal*, 47(4), 825–836.

Reniers, R. L., R. Corcoran, R. Drake, N. M. Shryane, and B. A. Völlm (2011): "The QCAE: A questionnaire of cognitive and affective empathy," *Journal of Personality Assessment*, 93(1), 84–95.

Reniers, R. L., R. Corcoran, B. A. Völlm, A. Mashru, R. Howard, and P. F. Liddle (2012): "Moral decision-making, ToM, empathy and the default mode network," *Biological Psychology*, 90(3), 202–210.

Sabini, J., M. Siepmann, J. Stein, and M. Meyerowitz (2000): "Who is embarrassed by what?," *Cognition and Emotion*, 14(2), 213–240.

Schaumberg, R. L., and F. J. Flynn (2012): "Uneasy lies the head that wears the crown: the link between guilt proneness and leadership," *Journal of Personality and Social Psychology*, 103(2), 327.

Schlenker, B. R. (2008): "Integrity and character: Implications of principled and expedient ethical ideologies," *Journal of Social and Clinical Psychology*, 27(10), 1078–1125.

Schlenker, B. R., J. R. Chambers, and B. M. Le (2012): "Conservatives are happier than liberals, but why? Political ideology, personality, and life satisfaction," *Journal of Research in Personality*, 46(2), 127–146.

Scholz, J. T., and M. Lubell (1998): "Trust and taxpaying: Testing the heuristic approach to collective action," *American Journal of Political Science*, 42(2), 398–417.

Seara-Cardoso, A., H. Dolberg, C. Neumann, J. P. Roiser, and E. Viding (2013): "Empathy, morality and psychopathic traits in women," *Personality and Individual Differences*, 55(3), 328–333.

Shepperd, J. A., W. A. Miller, C. T. Smith, and J. Algina (2014): "Does religion offer worldviews that dissuade adolescent substance use?," *Psychology of Religion and Spirituality*, 6(4), 292–301.

Sierra, J. J., and M. R. Hyman (2008): "Ethical antecedents of cheating intentions: Evidence of mediation," *Journal of Academic Ethics*, 6(1), 51–66.

Singhapakdi, A., S. J. Vitell, and G. R. Franke (1999): "Antecedents, consequences, and mediating effects of perceived moral intensity and personal moral philosophies," *Journal of the Academy of Marketing Science*, 27(1), 19–36.

Smith, B., and F. Shen (2013): "We All Think It's Cheating, But We All Won't Report It: Insights into the Ethics of Marketing Students," *Journal for Advancement of Marketing Education*, 21(1), 27–37.

Stoeber, J., and H. Yang (2016): "Moral perfectionism and moral values, virtues, and judgments: Further investigations," *Personality and Individual Differences*, 88, 6–11.

Sullivan, D., M. J. Landau, I. F. Young, and S. A. Stewart (2014): "The dramaturgical perspective in relation to self and culture," *Journal of Personality and Social Psychology*, 107(5), 767–790.

Swenson, C. (1996): "Experimental market evidence on implicit taxes," Working Paper.

Tangney, J., R. Dearing, P. Wagner, and R. Gramzow (2000): "The test of self-conscious afffec-3 (TOSCA-3).," Unpublished instrument.

Thomas, K. D. (2015): "The Psychic Cost of Tax Evasion," *Boston College Law Review*, 56, 617–670.

Thurman, Q. C., C. S. John, and L. Riggs (1984): "Neutralization and Tax Evasion: How effective would a moral appeal be in improving compliance to tax laws?," *Law & Policy*, 6(3), 309–327.

Torgler, B. (2002): "Speaking to theorists and searching for facts: Tax morale and tax compliance in experiments," *Journal of Economic Surveys*, 16(5), 657–683.

Torgler, B., and F. Schneider (2007): "What shapes attitudes toward paying taxes? Evidence from multicultural European countries," *Social Science Quarterly*, 88(2), 443–470.

——— (2009): "The impact of tax morale and institutional quality on the shadow economy," *Journal of Economic Psychology*, 30(2), 228–245.

Wahl, I., S. Muehlbacher, and E. Kirchler (2010): "The impact of voting on tax payments," *Kyklos*, 63(1), 144–158.

Wartick, M. L., S. A. Madeo, and C. C. Vines (1999): "Reward dominance in tax-reporting experiments: The role of context," *Journal of the American Taxation Association*, 21(1), 20–31.

Wenzel, M. (2005): "Motivation or rationalisation? Causal relations between ethics, norms and tax compliance," *Journal of Economic Psychology*, 26(4), 491–508.

Wolfe, R. N., R. D. Lennox, and B. L. Cutler (1986): "Getting along and getting ahead: Empirical support for a theory of protective and acquisitive self-presentation," *Journal of Personality and Social Psychology*, 50(2), 356–361.

Wowra, S. A. (2007): "Moral identities, social anxiety, and academic dishonesty among American college students," *Ethics & Behavior*, 17(3), 303–321.

Yoder, K. J., and J. Decety (2014): "Spatiotemporal neural dynamics of moral judgment: a high-density ERP study," *Neuropsychologia*, 60, 39–45.

# Appendix

## A Description of the questionnaires used in Experiment 1

| Questionnaires | Sub-scales | Measures | Description | Nb of items |
|---|---|---|---|---|
| Concern for Appropriateness Scale (CAS) | Cross-Situational Variability of Behavior (CSV) | Behavioral variability | CSV taps the behavioral variability that is a consequence of continually tailoring one's actions so as to avoid disapproval | 7 |
| | Attention to Social Comparison Information (ATSCI) | Tendency to compare behavior | Many of the items of the ATSCI subscale have a defensive connotation and of behavior comparison | 13 |
| Questionnaire for Cognitive and Affective Empathy (QCAE) | Perspective Taking (PT) | Intuitive perspective taking | Intuitively putting oneself in another person's shoes in order to see things from his/her perspective | 10 |
| | Online Simulation (OS) | Costly perspective taking | An effortful attempt to put oneself in another person's position by imagining what that person is feeling. Online simulation is likely to be used for future intentions | 9 |
| | Emotion Contagion (EC) | Emotional contagion | The automatic mirroring of the feelings of others | 4 |
| | Peripheral Responsivity (PERIR) | Mood transmission in a detached social context | The affective response when witnessing the mood of others in a detached social context | 4 |
| | Proximal Responsivity (PROXR) | Mood transmission in a closed social context | The affective response when witnessing the mood of others in a close social context | 4 |
| Guilt And Shame Proneness (GASP) | Guilt - Negative Behavior-Evaluations (NBE) | Guilt | Guilt - NBE items describe feeling bad about how one acted | 4 |
| | Guilt - Repair responses (GR) | Correction of transgressions | Guilt - repair items describe action tendencies (i.e., behavior or behavioral intentions) focused on correcting or compensating for the transgression | 4 |
| | Shame - Negative Self-Evaluations (NSE) | Shame | Shame - NSE items describe feeling bad about oneself | 4 |
| | Shame - Withdrawal Responses (SW) | Willingness to hide oneself | Shame - withdraw items describe action tendencies focused on hiding or withdrawing from public | 4 |

# B Description of the questionnaires used in Experiment 2

| Questionnaire | Sub-scales | Measures | Description | Nb of items |
|---|---|---|---|---|
| Ethics Position Questionnaire (EPQ) | Relativism | The extent to which the individual rejects universal moral rules in favor of relativism | Some individuals use moral absolutes in making moral judgments. Others do not rely on such universal moral rules | 10 |
| | Idealism | The extent to which the individual idealizes moral rules | Some individuals believe that "right" actions will always produce beneficial consequences. Others think that beneficial consequences will be mixed with non-beneficial ones | 10 |
| Integrity scale (IS) | Integrity | Level of integrity | Higher integrity involves personal commitment to moral identity that increases positive activities and reduces illicit temptations | 18 |
| Moralization of Everyday Life Scale (MELS) | Factor 1 | Use of deception | Lying or cheating to get something that one should not have had in the first place | 5 |
| | Factor 2 | Social norm violations that harm community members | Violations of social norm that harm community members | 5 |
| | Factor 3 | Laziness | Behaviors adopted because of laziness | 5 |
| | Factor 4 | Failures to do good | Failures to take an opportunity to act in a good way for the community | 5 |
| | Factor 5 | Violations of the body | Use or modifications of the body in ways that threaten body purity | 5 |
| | Factor 6 | Disgusting behaviors | Behaviors that are related to animal-like aspect of human nature | 5 |

# C Additional statistics on Experiment 1

Figure 3: Declared and earned income in the first experiment

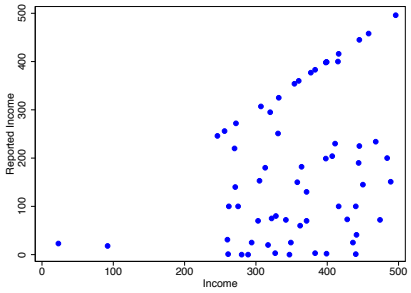


Table 9: Correlation matrix of the variables from the first experiment

| | CAS | CSV | ATSCI | COG. E. | PT | OS | AFF. E. | EC | PROXR | PERIR | GUILT | NBE | GR | SHAME | NSE | SW |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CAS | 1.0000 | | | | | | | | | | | | | | | |
| CSV | 0.7661 | 1.0000 | | | | | | | | | | | | | | |
| ATSCI | 0.8581 | 0.3275 | 1.0000 | | | | | | | | | | | | | |
| COG.E. | -0.0550 | 0.0150 | -0.0928 | 1.0000 | | | | | | | | | | | | |
| PT | -0.0708 | 0.0464 | -0.1412 | 0.7998 | 1.0000 | | | | | | | | | | | |
| OS | -0.0108 | -0.0273 | 0.0060 | 0.7411 | 0.1898 | 1.0000 | | | | | | | | | | |
| AFF.E | 0.1790 | 0.0810 | 0.1984 | 0.2068 | 0.0312 | 0.3034 | 1.0000 | | | | | | | | | |
| EC | 0.2417 | 0.1751 | 0.2155 | 0.0115 | -0.1057 | 0.1371 | 0.7702 | 1.0000 | | | | | | | | |
| PROXR | 0.1006 | -0.0103 | 0.1561 | 0.2688 | 0.0250 | 0.4118 | 0.8551 | 0.5803 | 1.0000 | | | | | | | |
| PERIR | 0.0631 | 0.0154 | 0.0805 | 0.2015 | 0.1587 | 0.1520 | 0.6688 | 0.1662 | 0.3852 | 1.0000 | | | | | | |
| GUILT | -0.1440 | -0.1502 | -0.0917 | 0.2423 | -0.0072 | 0.4045 | 0.3177 | 0.1749 | 0.4731 | 0.0844 | 1.0000 | | | | | |
| NBE | -0.1199 | -0.1013 | -0.0953 | 0.3134 | 0.0756 | 0.4281 | 0.3527 | 0.2325 | 0.4938 | 0.0850 | 0.9079 | 1.0000 | | | | |
| GR | -0.1223 | -0.1666 | -0.0467 | 0.0222 | -0.1402 | 0.1931 | 0.1246 | 0.0057 | 0.2374 | 0.0475 | 0.7248 | 0.3693 | 1.0000 | | | |
| SHAME | 0.4649 | 0.3897 | 0.3722 | 0.1273 | -0.0370 | 0.2496 | 0.3782 | 0.3078 | 0.3564 | 0.2028 | 0.3103 | 0.3716 | 0.0772 | 1.0000 | | |
| NSE | 0.3569 | 0.2892 | 0.2936 | 0.2074 | 0.0329 | 0.3026 | 0.3958 | 0.3412 | 0.3878 | 0.1773 | 0.4655 | 0.4789 | 0.2448 | 0.8754 | 1.0000 | |
| SW | 0.4292 | 0.3724 | 0.3334 | -0.0231 | -0.1127 | 0.0882 | 0.2140 | 0.1497 | 0.1826 | 0.1594 | -0.0047 | 0.0950 | -0.1666 | 0.7851 | 0.3878 | 1.0000 |

# D Additional statistics on Experiment 2

Figure 4: Earned and Declared (for WWF) income in the second experiment

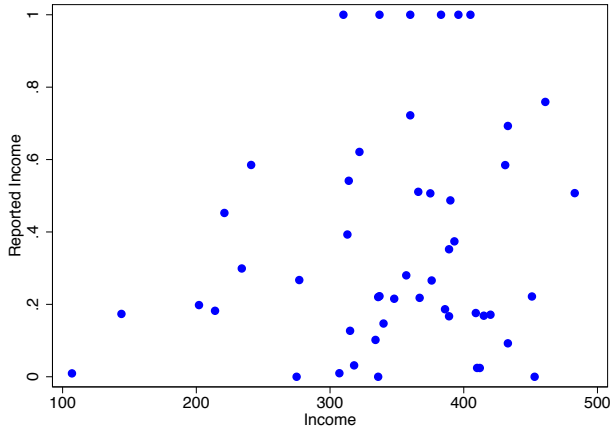



Table 10: Correlation matrix of the variables from the second experiment

| | Idealism | Relativism | Integrity | F1 | F2 | F3 | F4 | F5 | F6 |
|---|---|---|---|---|---|---|---|---|---|
| Idealism | 1.0000 | | | | | | | | |
| Relativism | 0.1435 | 1.0000 | | | | | | | |
| Integrity | 0.5995 | -0.1776 | 1.0000 | | | | | | |
| F1-Deception | 0.2980 | -0.2061 | 0.3747 | 1.0000 | | | | | |
| F2-Norm violation | 0.1388 | -0.0981 | -0.0282 | 0.4877 | 1.0000 | | | | |
| F3-Laziness | 0.0801 | 0.1460 | 0.1056 | 0.2002 | -0.0621 | 1.0000 | | | |
| F4-Failure | 0.2083 | -0.2284 | 0.2483 | 0.6572 | 0.5686 | 0.0885 | 1.0000 | | |
| F5-Body violations | 0.1517 | -0.0620 | 0.2501 | 0.5484 | 0.2796 | 0.6287 | 0.5442 | 1.0000 | |
| F6-Disgust | 0.0004 | -0.0625 | 0.0338 | 0.2597 | 0.5043 | 0.4143 | 0.2562 | 0.4361 | 1.0000 |

Table 11: Multivariate regressions of compliance (for ASPAS) decisions on psychometric scores

| **Variable** | **Coef.** | **(St. e.)** |
|---|---|---|
| _Idealism | 0.003 | (0.005) |
| _Relativism | -0.002 | (0.004) |
| _Integrity | -0.010 | (0.008) |
| _F1-Deception | 0.008 | (0.011) |
| _F2-Norm violation | 0.002 | (0.012) |
| _F3-Laziness | -0.005 | (0.015) |
| _F4-Failure | -0.001 | (0.011) |
| _F5-Body violations | 0.005 | (0.012) |
| _F6-Disgust | -0.007 | (0.009) |
| Intercept | 0.827 | (0.506) |

| | |
|---|---|
| N | 50 |
| $R^2$ | 0.09 |
| $F_{(9,40)}$ | .437 |



View publication stats