Journal of Economic Psychology xxx (2010) xxx–xxx

Contents lists available at ScienceDirect

Journal of Economic Psychology

journal homepage: www.elsevier.com/locate/joep

Sex differences in tax compliance: Differentiating between demographic sex, gender-role orientation, and prenatal masculinization (2D:4D)

Barbara Kastlunger [a,*,1], Stefan G. Dressler [a,1], Erich Kirchler [a], Luigi Mittone [b,2], Martin Voracek [a]

[a] *Faculty of Psychology, University of Vienna, Universitaetsstrasse 7, A-1010 Vienna, Austria*
[b] *CEEL – Computable and Experimental Economics Laboratory, University of Trento, Via Inama 5, I-38100 Trento, Italy*

ARTICLE INFO

Article history:
Received 27 February 2009
Received in revised form 18 March 2010
Accepted 22 March 2010
Available online xxxx

JEL classification:
H26
J16

PsycINFO classification:
3000
4200
2970

Keywords:
Tax compliance
Sex differences
Gender-role orientation
Digit ratio (2D:4D)
Prenatal testosterone

ABSTRACT

We used decision-making experiments to investigate tax compliance of women and men and focused on gender-role orientation as well as on the second-to-fourth digit ratio (2D:4D), a putative marker of prenatal testosterone exposure. In 60 experimental periods, participants were endowed with a certain amount of money representing income and had to pay taxes. They were audited with a certain probability and fined in case of detected evasion. Both demographic sex and gender-role orientation were significantly related to tax compliance, whereas 2D:4D was not. Women and less male-typical individuals were more compliant than men and more male-typical individuals. Women and men also differed regarding their taxpaying strategies. Whereas for men audits increased subsequent evasion, women's tax payments were less affected by prior audits.

1. Introduction

Differences between women and men in terms of their financial decisions are of continuing interest in research. Women were found to be more risk averse, especially in certain economic domains (Croson & Gneezy, 2009), and less likely to engage in unethical business behavior than men (Betz, O'Connell, & Shepard, 1989). Research on tax behavior suggests that women are more likely to cooperate and to evade to a lesser extent than men (Hasseldine, 1999). In this paper we use economic decision-making experiments to study tax compliance by women and men.

* Corresponding author. Tel.: +43 1427747808; fax: +43 1427747889.
E-mail addresses: barbara.kastlunger@univie.ac.at (B. Kastlunger), stefan.dressler@univie.ac.at (S.G. Dressler), erich.kirchler@univie.ac.at (E. Kirchler), luigi.mittone@economia.unitn.it (L. Mittone), martin.voracek@univie.ac.at (M. Voracek).
[1] Co-authors appear in alphabetical order.
[2] Tel.: +39 0461282213; fax: +39 0461282222.

0167-4870/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.joep.2010.03.015

There are two ways to approach the issue of tax compliance. There is the economic approach based on rational choice theory (Allingham & Sandmo, 1972) where taxes are paid or evaded strategically. Furthermore, there is a socio-psychological approach, where tax compliance is seen as determined by psychological factors like fairness considerations and moral perceptions (e.g., Kirchler, 2007).

In line with the literature on sex differences in risk-taking behavior (Croson & Gneezy, 2009) and tax compliance (Hasseldine, 1999), we assumed that women should exhibit higher levels of tax compliance, whereas men should show greater levels of evasion. The greater tendency among men to evade taxes was assumed to be related to (i) socialization differences that manifest themselves in different gender-role orientation and self-concepts of women and men and/or to (ii) underlying biological differences (foremost, prenatal testosterone exposure). We differentiated between these two indicators of gender vs. sex differences, i.e., gender vs. a biological indicator of masculinity (2D:4D). Research on gender differences in financial decision-making showed that apart from demographic sex the identification with masculine attributes also significantly influenced risk-taking (Meier-Pesti & Penz, 2008). To the best of our knowledge, 2D:4D as a biological marker for femininity/masculinity has not yet been studied in tax compliance research. Sapienza, Zingales, and Maestripieri (2009) suggested biological reasons for gender differences in financial risk aversion and in risky career choices by taking into account circulating testosterone and prenatal testosterone exposure. The research question we addressed regards differences between women and men in tax compliance and possible biological and/or social sources as causes of these differences.

The remainder of this paper is organized as follows: First, empirical results on demographic sex and gender in tax research and adjacent research areas are presented. Second, the methods and procedures of this study are described. Third, the effect of demographic sex on tax compliance is analyzed by differentiation between gender and a biological marker of early masculinization (2D:4D). Fourth, the results are discussed.

1.1. Differences between women and men in tax compliance

Tax compliance has been found to be related to perceived probability of audits and fines, as suggested by the rational choice model in economics (Allingham & Sandmo, 1972). Additional determinants were detected as well, however, such as age, sex, income, and psychological determinants (for a review, see Kirchler, 2007). Whereas empirical studies consistently support a positive relationship between age and tax compliance (e.g., Vogel, 1974), the findings regarding sex are less consistent (for an early review, see Jackson & Milliron, 1986). In the case of significant sex differences, women have often been found to be more compliant than men (e.g., Bazart & Pickhardt, 2009; Gerxhani, 2007; Hasseldine, 1999; Lewis, Carrera, Cullis, & Jones, 2009; Mason & Calvin, 1978; Porcano, 1988; Spicer & Becker, 1980; Spicer & Hero, 1985; Vogel, 1974; Wilson & Sheffrin, 2005). Several studies, however, considered sex differences as a side-effect rather than the central research question. One notable example of an exception was a study by Hasseldine (1999), who explicitly studied differences between women and men in tax compliance. Using self-reported responses from a sample of 605 taxpayers, he found differences between women and men in both attitudes towards non-compliance as well as in non-compliance behavior. Women were found to be more compliant than men.

Some studies failed to find sex differences or observed that women were not generally more compliant then men. For instance, Kirchler and Maciejovsky (2001) asserted that women's self-reported tax compliance was lower than men's. Friedland, Maital, and Rutenberg (1978) found women more likely to evade taxes, but to a smaller extent than men. Chung and Trivedi (2003) found women more compliant than men only after providing the sample with persuasive reasons to pay taxes. Wenzel (2002) found women to be more compliant with regard to reported income and deduction claims, but no sex differences regarding reports of extra income. Torgler and Schneider (2004) reported higher tax morale for women than men in Switzerland and Belgium, but minor differences in Spain. It has to be noted that results not indicating differences between women and men are often not even reported (Unger, 1979). Although empirical findings are mixed, when there were significant sex differences, women were more compliant than men.

Differences in tax compliance between women and men might be because of different ethical standards (Chung & Trivedi, 2003; Grasso & Kaplan, 1998; Torgler & Valev, 2006), as well as due to differences in risk propensity (for a general review, see Byrnes, Miller, & Schafer, 1999). Compared with men, women were found to overestimate the detection probability and severity of fines in case of detection (Hasseldine, 1999; Kinsey, 1992; Richards & Tittle, 1981; Smith, 1992). Survey data suggest that differences between women and men in financial risk-taking interact with other demographic variables, like marital status (Jianakoplos & Bernasek, 1998; Sunden & Surette, 1998).

Most of these studies on differences between women and men considered only demographic sex. Differences between women and men may be caused, not only by biological differences, but also by gender-role orientation and characteristics associated with gender (feminine traits, such as socially desirable behavior and friendliness; and masculine traits, such as dominance, competitiveness, and aggression; see Bem, 1981; Eagly, 1987). Differences between women and men in tax compliance should be interpreted with caution because effects might be owed to factors covarying with demographic sex, such as socialization and education, or with interest and involvement in the topic at stake.

A differentiated usage of the terms "gender" and "sex" is needed rather than inappropriate use of the two concepts as synonymous (Gentile, 1998). Whereas "gender" refers to cultural influence, social categorization, and identity, "sex" is related to characteristics that are caused by underlying biological differences (Anselmi & Law, 1998; Unger, 1979).

Please cite this article in press as: Kastlunger, B., et al. Sex differences in tax compliance: Differentiating between demographic sex, gender-role orientation, and prenatal masculinization (2D:4D). *Journal of Economic Psychology* (2010), doi:10.1016/j.joep.2010.03.015

1.2. Gender-role orientation

Anselmi and Law (1998) defined gender as a ''social category of shared meanings about characteristics of maleness and femaleness and the behaviors, attitudes, and feelings associated with those characteristics. How we view ourselves and how others treat us will often be directly influenced by what position in the social category of gender we reside'' (p. 2). Women and men adopt specific culturally determined roles in our society. Gender roles include descriptions of typical women and typical men as well as prescriptions in the sense of expectations about traits and behaviors (Alfermann, 1996).

The differentiation between gender or gender-role orientation and biological sex has rarely been analyzed in economic decision-making. In financial risk-taking, a study by Meier-Pesti and Penz (2008) addressed the puzzle of sex and gender-role orientation. Women and men indicated their investment preferences and attitudes towards risk. Furthermore, sex-role identity was assessed. In contrast to femininity, masculinity was found to be positively related to risk-taking in investment decisions. Differences between women and men in financial risk-taking decreased with controls for femininity and masculinity. The researchers stated that differences between women and men in financial risk-taking might be owed to different levels of identification with masculine characteristics. They supposed, however, that owing to changes in social roles women steadily adopt more and more masculine characteristics which might lead to assimilation of women's and men's financial behavior over time.

1.3. 2D:4D: Biological marker of prenatal masculinization

Biological causes of differences between women and men are, inter alia, ascribed to sex-hormonal factors. For example, high testosterone in men is related to rejection of unfair offers in ultimatum games (Burnham, 2007) and to dominant and antisocial behavior in order to maintain high status (for a review, see Mazur & Booth, 1998).

It is not only circulating sex hormones that might influence behavior, however, but also exposure to sex hormones in prenatal life. According to Hines (2004), ''infants enter the world with some predispositions to 'masculinity' and 'femininity', and these predispositions appear to result largely from sex hormones to which they were exposed before birth'' (p. 10). The degree of prenatal masculinization is now frequently assessed with the digit ratio (2D:4D), i.e., the length ratio of the second (2D) to the fourth finger (4D). 2D:4D is widely believed to be a biomarker for prenatal testosterone levels and thus a retrospective indicator of masculinization when measured in adults (for a review, see Manning, 2002). It is supposed to ''provide a simple and widely-available method for examining hormonal effects on human behavior'' (Cohen-Bendahan, van de Beek, & Berenbaum, 2005, p. 373). Lower 2D:4D indicates higher prenatal testosterone levels. On average, men's 2D:4D is lower than women's. (Manning, 2002; Voracek, Manning, & Dressler, 2007). Sex differences in finger length patterns have been found consistently across different world regions (Peters, Tan, Kang, Teixeira, & Mandal, 2002).

So far more than 300 reports in 2D:4D research have been published (Voracek & Loibl, 2009). Various types of economic behaviors have been found to be related to 2D:4D. For instance, a masculine (lower) 2D:4D was found to be associated with preference for fair and cooperative (and not extensively altruistic or extremely egoistic) behavior in social dilemma situations (Millet & Dewitte, 2006). Coates, Gurnell, and Rustichini (2009) analyzed 2D:4D in financial traders and found that lower 2D:4D predicted traders' long-term profitability.[3] Using a sample of 550 MBA students, Sapienza et al. (2009) provided evidence of the activational (salivary testosterone) and organizational effects of testosterone (2D:4D, performance in the Baron–Cohen test) on risk aversion and thus students' career choices. The authors found that salivary testosterone was negatively correlated with risk aversion in women but not in men. Sex differences disappeared, however, when they analyzed women and men with comparable low salivary testosterone levels, indicating a nonlinear relation between testosterone and risk aversion. Although the expected negative effect of 2D:4D on the probability of starting a risky career in finance was found, the effect of 2D:4D on risk aversion was low and again mostly driven by women. In an investment game conducted by Apicella et al. (2008) financial risk-taking was positively related to circulating testosterone (assessed via saliva samples) and to pubertal testosterone exposure (measured via facial masculinity), whereas, risk-taking was not related to prenatal testosterone exposure as assessed via 2D:4D. These authors attributed the null finding to ethnic differences in the sample and to the small sample size. The inconsistent results on 2D:4D were enriched by studies claiming to analyze interactions of 2D:4D with situational cues (e.g., Millet & Dewitte, 2008, 2009; Ronay & von Hippel, 2009; van den Bergh & Dewitte, 2006). For instance 2D:4D was found to interact with sexual cues on decisions in ultimatum games (van den Bergh & Dewitte, 2006), with situational aggression cues on pro-social behavior in dictator games (Millet & Dewitte, 2009), with induced status on monetary discounting (Millet & Dewitte, 2008) and with power in risk-taking behavior (Ronay & von Hippel, 2009).

There is also some evidence for associations of 2D:4D and gender-role orientation. Csathó et al. (2003) found lower 2D:4D to be associated with self-reported masculine sex-role identity in Hungarian women. Studies on 2D:4D and gender-role orientation are still rare, however, and findings have been inconsistent (Lippa, 2006; Rammsayer & Troche, 2007; Schmukle, Liesenfeld, Back, & Egloff, 2007).

In the present study, we investigated tax compliance and strategic behavior of women and men. Whereas previous studies considered demographic sex as a determinant of financial decision-making, without paying much attention to underlying

[3] The relation between 2D:4D and success in high-frequency trading was subsequently discussed in the light of need for achievement (see letters by Coates, 2009; Millet, 2009).

Please cite this article in press as: Kastlunger, B., et al. Sex differences in tax compliance: Differentiating between demographic sex, gender-role orientation, and prenatal masculinization (2D:4D). *Journal of Economic Psychology* (2010), doi:10.1016/j.joep.2010.03.015

factors (with the exception of Meier-Pesti & Penz, 2008), we additionally focused on the impact of gender-role orientation owing to socialization and presumed biological influences (2D:4D).

2. Method

2.1. Participants

One hundred and sixteen young adults were recruited through announcements on the bulletin board of a Faculty of Economics at a University in northern Italy. Data of nine participants were incomplete, leaving a valid sample for analysis comprised of 43 female and 64 male Italian students, aged between 20 and 35 years ($M = 23.8$, $SD = 2.9$).

2.2. Experimental design and procedures

Sessions were run with groups of 11–15 participants in a computerized laboratory. The experimental design was a repeated-measures design and consisted of 60 periods (see Kastlunger, Kirchler, Mittone, & Pitters, 2009). In order to guarantee anonymity, participants were given an identification code which they had to enter in their computer file. Participants were told that the experiment simulated a real taxpaying context. They were instructed to file their tax returns in each period after being provided with a constant income of 1000 ECU (Experimental Currency Unit). They were informed that the tax rate was 20%, and that in the case of detection they had to refund the evaded taxes and pay a fine in relation to their evasion. The probability of being audited was 15%. After reading the instructions, participants filed taxes over 60 taxpaying periods on their computer. Audits occurred after periods 2, 3, 7, 9, 14, 18, 20, 31, and 51. Participants were neither informed of the exact number of periods nor of the number of audits. The dominant strategy to maximize individual profits was a full tax evasion in all periods.

After the tax filing periods, items on self-reported taxpaying behavior (see Section 2.3.1.) as well as questionnaires on gender-role orientation (see Section 2.3.3.) were presented. After the experimental sessions, participants were paid according to their performance in the experiment and flatbed-scanned image files of their hands were taken for measurement of 2D:4D (see Section 2.3.2.).

2.3. Material

2.3.1. Self-reported taxpaying behavior

Participants answered six items on their perceptions regarding the audits and their strategic behavior (e.g.,"The probability to be audited was very low"; "I used a special strategy when deciding how many taxes to pay."). The items were presented with a five-point scale (presented in Table 4).

2.3.2. Biological marker of masculinity (2D:4D)

Following standard practice of 2D:4D research (Voracek & Dressler, 2006; Voracek et al., 2007), the length of participants' index (2D) and ring (4D) fingers was measured twice[4] using hand-scans with a gap of two days between the measurement. Intraclass correlation coefficients (*ICC*)[5] were as follows (all $ps < .001$, with $df_1 = 106$ and $df_2 = 106$ for the F ratios): for the right index finger, $ICC = .998$ ($F = 983.24$); for the right ring finger, $ICC = .997$ ($F = 710.16$); for the left index finger, $ICC = .999$ ($F = 1815.79$); and for the left ring finger, $ICC = .999$ ($F = 2261.69$); for right-hand 2D:4D, $ICC = .975$ ($F = 78.51$); and for left-hand 2D:4D, $ICC = .992$ ($F = 240.33$). In accordance with the literature (Manning, 2002; Voracek et al., 2007), we found 2D:4D to be significantly higher for women than for men, with this sex effect being statistically significant. Descriptive statistics and results for tests for sex differences are presented in Table 1.

2.3.3. Gender-role orientation

Gender-role orientation questionnaires consisted of the femininity (PAQ/F) and the masculinity (PAQ/M) scales of the Italian version of the Personal Attributes Questionnaire (PAQ; Bonnes-Dobrowolny & Vicarelli, 1982), a 50-item preference rating scale of female-typical and male-typical occupations (gender diagnosticity – occupational preferences – GD_OP; Lippa, 2002), and the six items of the Sex Role Identity Scale (SRIS; Storms, 1979).[6]

[4] Measurements were made blind to the other study data collected from high-resolution laser printouts of the flatbed-scanned images of right and left hands by an experienced investigator using a digital vernier caliper measuring to 0.01 mm (Mitutoyo Ltd., Andover, Hampshire, U.K.; Model 500-191U). Measurement landmarks were the ventrally located proximal-most (boundary) metacarpophalangeal flexion crease that divides the finger from the palm region and the finger tip. 2D:4D was calculated by dividing the length of the index finger by the length of the ring finger (right hand: R2D:4D; left hand: L2D:4D).

[5] ICC (two way mixed-effects model and absolute-agreement definition, see McGraw & Wong, 1996; Voracek & Dressler, 2006) of finger-length measures were used to quantify the reproducibility of the two measurements.

[6] A similar combination of psychometric measures of gender was used by Lippa (2002).

Table 1
Correlations of demographic sex, 2D:4D, gender-related scales, gender-role orientation and average tax payments.

		M(*SD*)	*t*(*df* = 104)	R2D:4D	L2D:4D	PAQ/F	PAQ/M	SRIS	GD_OP	Gender-role orientation	Average tax payments
Demographic sex				−.33**	−.28**	−.37**	.35**	.87**	.76**	.85**	−.27**
R2D:4D	Women Men	0.975(0.031) 0.953(0.030)	3.62**	0.962 (0.032)	.80**	.15	−.12	−.30**	−.24*	−.29**	.07
L2D:4D	Women Men	0.975(0.031) 0.956(0.034)	2.93**	.79** .77**	0.963 (0.034)	.12	−.02	−.22*	−.17^{+}	−.20*	.05
PAQ/F (α = .78)	Women Men	4.15(0.44) 3.77(0.50)	4.00**	.09 .01	−.04 .06	3.92 (0.51)	−.03	−.36**	−.39**	−.58**	.21*
PAQ/M (α = .72)	Women Men	3.47(0.46) 3.81(0.45)	−3.79**	.05 −.05	.06 .10	.31* .00	3.68 (0.48)	.34**	.36**	.54**	−.12
SRIS (α = .96)	Women Men	1.98(0.60) 4.10(0.58)	−18.10**	−.36* .22^{+}	−.30^{+} .24^{+}	−.35* .05	.04 .12	3.26 (1.20)	.68**	.86**	−.27**
GD_OP (α = .78)	Women Men	2.14(0.69) 3.61(0.59)	−11.78**	.00 .04	−.06 .16	−.05 −.28	.16 .15	−.13 .20	3.03 (0.96)	.88**	−.27**
Gender-role orientation	Women Men	−1.05(0.51) 0.68(0.56)	−16.25**	−.16 .07	−.11 .17	−.47** −.58**	.45** .53**	.50** .48**	.63** .72**	0.00 (1.00)	−.31**
Average tax payments	Women Men	139.70(52.42) 111.08(48.95)	2.86**	.05 −.08	.06 −.09	.05 .17	.04 −.08	−.32* .10	.04 −.21	−.11 −.17	122.42 (52.04)

Note: N = 106. Diagonal entries are the *M* and *SD* values for the total sample; correlations for the whole sample are presented in the right upper corner; correlations for women (n = 42) and men (n = 64) are shown separately in the lower left corner.
$^{+}$ $p < .10$.
* $p < .05$.
** $p < .01$.

PAQ. Continuous PAQ/F and PAQ/M scores (high values indicating high femininity in the PAQ/F and high masculinity in the PAQ/M, respectively) were computed. Moreover, participants were categorized by median-splitting the PAQ/F (Md = 3.92) and PAQ/M (Md = 3.69) into masculine (high PAQ/M, low PAQ/F; n = 28), feminine (low PAQ/M, high PAQ/F, n = 26), androgen (high PAQ/M, high PAQ/F, n = 28) and undifferentiated (low PAQ/M, low PAQ/F, n = 25) type (see Bonnes-Dobrowolny & Vicarelli, 1982, p. 131).

GD_OP. In accordance with Lippa, gender-diagnostic probabilities of being male were computed by applying five discriminant analyses with stepwise selection of variables (for details, see Lippa, 1991, 1995; Lippa & Connelly, 1990) by assuming prior probabilities based on group size. The five resulting probabilities of being male were summed up to one GD_OP score (high values indicate high preference for masculinity-related occupations).

SRIS. Owing to the high correlation between the masculinity and the femininity items of the SRIS ($r = -.80$, $p < .01$), the three items measuring femininity were inverted in order to build one SRIS score (high values indicate high masculinity).

Reliabilities of the three psychometric measures of gender were acceptable and means of each scale differed between women and men in the expected direction: women described themselves as more feminine and less masculine, preferred less male-typed occupations, and scored lower on masculine sex-role identity. Men showed the opposite pattern. Sample reliabilities of these scales and descriptive statistics are displayed in Table 1.

2.4. Preliminary analyses

Although Lippa (1991), Lippa and Connelly (1990) reported factorial distinctiveness of gender diagnosticity from the PAQ, we decided to aggregate the three psychometric measures of gender utilized in this study into one gender-role orientation score in order to have one overall indicator of masculine gender-role identity composed of scales of masculinity with different theoretical approaches. A factor analysis of mean item responses on the four scales (PAQ/F, PAQ/M, GD_OP, SRIS) was conducted, using principal component method and constraint of one factor to be extracted (eigenvalue γ = 2.14, accounting for 53.5% of the variance). The resulting factor values were used as a gender-role orientation score, with higher values indicating higher masculinity (range: −2.26 to 1.80).[7]

[7] For all independent variables (i.e. demographic sex, gender-role orientation, and R2D:4D) and average tax compliance, Mahalanobis distance was calculated to identify outliers. One participant was identified as an outlier (D^2 = 13.37, p = .01; demographic sex = female, R2D:4D = 0.902, gender-role orientation = 0.25, average tax payments = 200 ECU) and was therefore excluded from further analyses.

3. Results

3.1. Demographic sex, 2D:4D, gender-role orientation, and tax compliance

In this section we analyzed Pearson correlations between demographic sex, 2D:4D, psychometric measures of gender (PAQ/F, PAQ/M, SRIS, GD_OP) as well as overall gender-role orientation and average tax payments over 60 periods for the whole sample and separately for women and men (Table 1).

Demographic sex was significantly related with each of the other indicators. Being a man was related to lower 2D:4D (for R2D:4D: $r = -.33$, $p < .01$; for L2D:4D: $r = -.28$, $p < .01$) and to more masculine self-descriptions on the psychometric measures of gender (PAQ/F: $r = -.37$, $p < .01$; PAQ/M: $r = .35$, $p < .01$; SRIS: $r = .87$, $p < .01$; GD_OP: $r = .76$, $p < .01$). Accordingly, men scored higher on the gender-role orientation score ($r = .85$, $p < .01$), indicating higher masculinity.

2D:4D. Participants' right-hand and left-hand 2D:4D were strongly correlated within the sexes (for women: $r = .79$, $p < .01$; for men: $r = .77$, $p < .01$). Analyzing the relations between the psychometric measures of gender and 2D:4D, we found that females' sex-role identity (SRIS) related strongly to their right-hand 2D:4D ($r = -.36$, $p < .05$) and only marginally to their left-hand 2D:4D ($r = -.30$, $p = .06$). Women with higher (more feminine) 2D:4D described themselves as less masculine in the SRIS, whereas women with lower (more masculine) 2D:4D described themselves as more masculine. No significant correlations were found between the general gender-role orientation score and 2D:4D (women, R2D:4D: $r = -.16$, $p = .32$; and L2D:4D: $r = -.11$, $p = .50$; men, R2D:4D: $r = .07$, $p = .58$; and L2D:4D $r = .17$, $p = .17$).

Gender-role orientation. The psychometric measures of gender were significantly correlated with each other in the expected direction. Participants with higher masculine self-ascriptions (PAQ/M) indicated higher preferences for male-typical occupations (GD_OP: $r = .36$, $p < .01$) and scored higher on masculine sex-role identity (SRIS: $r = .34$, $p < .01$). Feminine self-ascriptions (PAQ/F) correlated negatively with masculine occupational preferences (GD_OP: $r = -.39$, $p < .01$) and with masculine sex-role identity (SRIS: $r = -.36$, $p < .01$). Masculine sex-role identity (SRIS) was positively correlated with masculine occupational preferences (GD_OP: $r = .68$, $p < .01$). In line with Bonnes-Dobrowolny and Vicarelli (1982), PAQ/F and PAQ/M were not correlated ($r = -.03$, $p = .80$), indicating that the two scores are independent of each other.

Average tax payments over 60 taxpaying periods were significantly related to demographic sex ($r = -.27$, $p < .01$), indicating a significant relation between being a man and evading taxes. Tax compliance was significantly related to femininity in the PAQ (PAQ/F: $r = .21$, $p = .03$), to sex-role identity (SRIS: $r = -.27$, $p < .01$), gender diagnosticity (GD_OP: $r = -.27$, $p < .01$), and to overall gender-role orientation ($r = -.31$, $p < .01$), indicating that participants with less feminine self-attributions, more masculine sex-role identity, more masculine occupational preferences, and a more masculine gender-role orientation in general paid less tax during the 60 taxpaying periods. Within the sexes, only one significant relation between women's average tax compliance and SRIS ($r = -.32$, $p = .04$) emerged. We did not find any significant correlation between 2D:4D and tax evasion (women: R2D:4D: $r = .05$, $p = .78$; L2D:4D: $r = .06$, $p = .72$; men: R2D:4D: $r = -.08$, $p = .55$; L2D:4D: $r = -.09$, $p = .50$).

3.2. Disentangling the effects of demographic sex, 2D:4D, and gender-role orientation on tax compliance

To analyze the relations between demographic sex and tax compliance and to differentiate between aspects of gender socialization and biological determinants of sex differentiation (2D:4D), we calculated ML random effects models for panel-structured data with STATA 10.[8]

Three regression models were conducted, by including step by step (i) demographic sex, (ii) R2D:4D and (iii) gender-role orientation (see Table 2). Only 2D:4D of participants' right-hands was used, because it is assumed to express early testosterone exposure more strongly than 2D:4D of the left-hand (Manning, 2002). We controlled for within-subject correlations over the 60 periods by including tax payments as a lagged variable and accounted for the immediate effect of audits by controlling for audits in the previous period.

In the first regression model, (i) demographic sex significantly influenced tax compliance (*Coef.* = −23.27, *SE* = 8.20, $p < .01$) indicating that men paid less taxes. In the second regression model, (ii) R2D:4D was introduced. The influence of this biological marker on tax compliance was not significant (*Coef.* = −35.57, *SE* = 133.96, $p = .79$), demographic sex remained significant (*Coef.* = −24.04, *SE* = 8.70, $p < .01$). In the third regression model (iii), gender-role orientation was introduced. No influence of gender-role orientation (*Coef.* = −11.55, *SE* = 7.45, $p = .12$) on tax payments was found and the effect of demographic sex (*Coef.* = −4.13, *SE* = 15.46, $p = .79$) disappeared, which might be explained by the strong correlation between these two variables (see Table 1).[9]

Due to possible problems of collinearity, we proceeded by using a dummy variable of the masculine-type of the PAQ-categorization (see Section 2.3.3.; masculine-type = 1, other types = 0) that, compared to the overall gender-role orientation, was less correlated with demographic sex ($r = .35$, $p < .01$), though it showed a similar correlation with average tax payments ($r = -.24$, $p = .01$). Therefore, we included the masculine-type dummy instead of the overall gender-role identity in a fourth

[8] We also calculated random effects probit models and found only weak effects of demographic sex on the probability of being compliant or not. The results are presented in Appendix A.

[9] Calculating a linear regression model with average tax payments as dependent variable variability inflation factors resulted as follows: demographic sex VIF = 3.65, gender-role identity VIF = 3.54; R2D:4D VIF = 1.13.

Please cite this article in press as: Kastlunger, B., et al. Sex differences in tax compliance: Differentiating between demographic sex, gender-role orientation, and prenatal masculinization (2D:4D). *Journal of Economic Psychology* (2010), doi:10.1016/j.joep.2010.03.015

Table 2
ML random effects model (MLE) for panel-structured data for tax payments on demographic sex, 2D:4D, and gender-role orientation.

Tax payments on	i		ii		iii		iv	
	Coef.	*SE*	*Coef.*	*SE*	*Coef.*	*SE*	*Coef.*	*SE*
Constant	119.06**	6.61	153.73	130.77	145.05	129.43	149.50	129.04
Tax payments in the previous period	0.18**	0.01	0.18**	0.01	0.18**	0.01	0.18**	0.01
Audit in the previous period (audit = 1)	−27.74**	2.46	−27.74**	2.46	−27.74**	2.46	−27.74**	2.46
Demographic sex (male = 1)	−23.27**	8.20	−24.04**	8.70	−4.13	15.46	−18.71*	9.13
R2D:4D			−35.57	133.96	−39.07	132.49	−30.03	132.21
Gender-role orientation (masculinity)					−11.55	7.45		
Masculine type (PAQ)							−16.32+	9.59
Sigma_u	40.23	2.98	40.21	2.97	39.74	2.94	39.64	2.94
Sigma_e	70.07	0.63	70.07	0.63	70.07	0.63	70.07	0.63
Rho	0.25	0.03	0.25	0.03	0.24	0.03	0.24	0.03

Note: $N = 106$.
+ $p < .10$.
* $p < .05$.
** $p < .01$.

regression model (iv). Being categorized as masculine-type revealed a trend of a significant influence on tax payments (*Coef.* = −16.32, *SE* = 9.59, *p* = .09). The influence of demographic sex was reduced (*Coef.* = −18.71, *SE* = 9.13, *p* = .04) and R2D:4D was still not significant (*Coef.* = −30.03, *SE* = 132.21, *p* = .82). Although, a mediator effect (Baron & Kenny, 1986) of masculine gender-role on demographic sex and tax compliance was not completely significant, trends were revealed. The results highlighted, however, that differences between women and men in tax compliance owed more to gender-role orientation than to prenatal testosterone exposure.

We further analyzed the influence of 2D:4D and gender-role orientation within sexes but found no significant effects. In accordance to context-dependency of 2D:4D we also analyzed its interaction with experienced audits but found no significant relation with tax contributions.

3.3. Demographic sex and taxpaying strategies

In order to analyze differences in taxpaying strategies between women and men, we calculated a ML random effects model, introducing the interaction between being a man (demographic sex) and being audited in the previous period. Results are presented in Table 3.

The significant interaction effect (*Coef.* = −32.48, *SE* = 5.02, $p < .01$) suggested that experiencing an audit affected women and men differently, in that men paid less tax after experiencing an audit compared with women. Men exploited strategically advantageous, but risky, situations in the taxpaying simulations more extensively than women did.

To support our results, we requested participants to report their perceptions on the audits and their strategic behavior after the experiment (see Table 4). Many more men indicated adopting special strategies when paying their taxes ("I used a special strategy when deciding how much tax to pay"; women: $M = 2.54$, $SD = 1.38$; men: $M = 3.36$, $SD = 1.44$; $t(103) = -2.90$; $p < .01$) and fewer of them gave up their initial strategy because of audits that they could not control ("I gave up using a strategy to increase my income because there was no way to control the audits"; women: $M = 3.05$, $SD = 1.20$; men: $M = 2.42$, $SD = 1.31$; $t(103) = 2.47$; $p = .02$).

Furthermore, gender-role orientation was significantly related to acting strategically. Participants with a more masculine gender-role orientation exhibited more strategic taxpaying behavior ($r = .29$, $p < .01$). Moreover, giving up one's strategy because of uncontrollable audits was negatively correlated with masculine gender-role identity ($r = -.26$, $p < .01$).

Table 3
ML random effects model (MLE) for panel-structured data for tax payments on demographic sex and the interaction between demographic sex and experienced audit.

Tax payments on	*Coef.*	*SE*
Constant	116.04**	6.63
Tax payments in the previous period	0.18**	0.01
Audit in the previous period (audit = 1)	−8.13*	3.90
Demographic sex (male = 1)	−18.31*	8.23
Demographic sex by audit in the previous period	−32.48**	5.02
Sigma_u	40.22	2.97
Sigma_e	69.83	0.63
Rho	0.25	0.03

Note: $N = 106$.
* $p < .05$.
** $p < .01$.

Table 4
Participants' self-reports on audits and strategic taxpaying behavior.

Item (1 = no agreement to 5 = strong agreement)	Total sample *M* (*SD*)	Women *M* (*SD*)	Men *M* (*SD*)	*t*(*df*)	*p*
The probability to be audited was very low.	3.09 (1.14)	2.95 (1.02)	3.17 (1.20)	−1.01(94.91)[#]	=.32
I had the impression of being "chased" by the financial office.	2.10 (1.21)	2.02 (1.08)	2.14 (1.30)	−0.48(103)	=.63
I used a special strategy when deciding how much tax to pay.	3.04 (1.47)	2.54 (1.38)	3.36 (1.44)	−2.90(103)	<.01
I had bad luck with the number of audits.	2.37 (1.20)	2.15 (1.01)	2.52 (1.29)	−1.64(98.46)[#]	=.11
I gave up using a strategy to increase my income because there was no way to control the audits.	2.67 (1.30)	3.05 (1.20)	2.42 (1.31)	2.47(103)	=.02
I had the impression of being audited more often than the other participants.	1.86 (1.16)	1.85 (1.06)	1.86 (1.22)	−0.03(103)	=.98

[#] Welch's test owing to unequal variance.

Analysis of 2D:4D and self-reported tax behavior revealed only one significant positive relationship between women's 2D:4D and the item "I used a special strategy when deciding how much tax to pay" (R2D:4D: $r = .28$, $p = .08$; L2D:4D: $r = .31$, $p = .05$). This finding loses significance after Bonferroni correction for multiple statistical testing and should therefore be interpreted with caution.

4. Discussion

The principal aim of this study was to investigate taxpayers' compliance and strategic taxpaying behavior, as related to demographic sex, gender-role orientation, and a marker of prenatal masculinization (2D:4D). Several tax compliance studies found women to be more compliant than men. The present study was explicitly designed to test both sex and gender differences in tax compliance and replicated and extended these findings. In addition, women and men differed with regard to tax compliance and also with regard to taxpaying strategies. Men were less compliant than women and were more likely to act strategically when paying taxes.

In the present study we tried to distinguish clearly between psychological, social, and biological sex with regard to tax behavior. Gender-role orientation was assessed by three psychometric measures of gender (Italian version of the PAQ, SRIS, and gender diagnosticity). Significant associations were found between gender-role orientation and tax compliance, in that participants scoring higher on masculinity were more likely to evade taxes to increase profit. Moreover, we focused on the question whether biological differences at the level of fetal sex hormones might also be related to tax compliance. Therefore, 2D:4D was used to elucidate the relationships between biological aspects of masculinization and tax compliance. Although 2D:4D was reliably measured and differed between sexes in the expected direction, we found no significant association between 2D:4D and average tax compliance. It may be argued that the sample size in our experiment was small and that at least the direction of the correlations among women was as expected.

The present findings are similar to those of Apicella et al. (2008), who also did not obtain evidence of significant relations between 2D:4D and financial risk-taking, as well as of those of Sapienza et al. (2009), who found only small effects for 2D:4D on risk aversion that did not reach significance. With reference to these two studies, it might be argued that the activational aspect of testosterone, measured for example via saliva samples, might be more adequate in predicting financial risk-taking behavior than the organizing effect of testosterone measured by 2D:4D. Moreover, recent findings suggest that 2D:4D may mainly be relevant in the presence of stimulating context cues (e.g., Millet & Dewitte, 2008, 2009; Ronay & von Hippel, 2009; van den Bergh & Dewitte, 2006). We controlled for context dependence of 2D:4D by analyzing its interaction with experienced audits, but did not find any significant effect on tax compliance. The interaction of 2D:4D and relevant situational cues in taxpaying decisions would be a fruitful area for future research.

The results of our study suggest that women's higher tax compliance compared with men's might be better explained by differences in socialization, self-image, and femininity–masculinity traits, rather than by prenatal hormonal differences, as measured by 2D:4D. If sex differences in tax compliance are more socially than biologically determined, such differences may well change over time with changing gender-specific socialization (cf. Meier-Pesti & Penz, 2008). It has to be noted, however, that it is difficult, if not impossible, to determine clearly whether biological or social causes determine differences between men and women in taxpaying behavior. Differences between the sexes are probably owed to an interaction of both (cf. Anselmi & Law, 1998; Unger & Crawford, 1998).

Acknowledgments

Special thanks are owed to Michael Wenzel, Martin Ejnar Hansen, and Julia Pitters for critical comments on this paper and to Marco Tecilla and his colleagues for technical and experimental support. This research was partly financed by the University of Trento, Italy, the University of Vienna, Austria, and a grant to Erich Kirchler from the Austrian Science Fund (FWF), Project Number AP1992511.

Table A1
Random effects probit model for tax compliance on demographic sex, 2D:4D, and gender-role orientation.

Tax compliance on	i		ii		iii		iv	
	Coef.	*SE*	*Coef.*	*SE*	*Coef.*	*SE*	*Coef.*	*SE*
Constant	0.05	0.15	1.19	3.00	1.01	2.98	1.11	2.99
Tax compliance in the previous period	0.35**	0.04	0.35**	0.04	0.35**	0.04	0.35**	0.04
Audit in the previous period (audit = 1)	−0.42**	0.05	−0.42**	0.05	−0.42**	0.05	−0.42**	0.05
Demographic sex (male = 1)	−0.33+	0.19	−0.35+	0.20	0.04	0.36	−0.28	0.21
R2D:4D			−1.17	3.08	−1.23	3.05	−1.07	3.07
Gender-role orientation (masculinity)					−0.23	0.17		
Masculine type (PAQ)							−0.23	0.22
/lnsig2u	−0.16	0.17	−0.16	0.17	−0.18	0.17	−0.17	0.17
Sigma_u	0.92	0.08	0.92	0.08	0.91	0.08	0.92	0.08
Rho	0.46	0.04	0.46	0.04	0.46	0.04	0.46	0.04
Number of observations (participants)	6254(106)		6254(106)		6254(106)		6254(106)	
Wald χ^2	147.55**		147.73**		149.78**		148.98**	

Note: N = 106.
+ $p < .10$.
* $p < .05$.
** $p < .01$.

Table A2
Random effects probit model for tax compliance on demographic sex and the interaction of demographic sex and experienced audit.

Tax compliance on	*Coef.*	*SE*
Constant	0.005	0.15
Tax compliance in the previous period	0.35**	0.04
Audit in the previous period (audit = 1)	−0.15+	0.08
Demographic sex (male = 1)	−0.26	0.19
Demographic sex by audit in the previous period	−0.44**	0.11
/lnsig2u	−0.16	0.17
Sigma_u	0.92	0.08
Rho	0.46	0.04
Number of observations (participants)	6254(106)	
Wald $\chi^2(4)$	161.92**	

Note: N = 106.
+ $p < .10$.
* $p < .05$.
** $p < .01$.

Appendix A

See Tables A1 and A2.

References

Alfermann, D. (1996). *Geschlechtsrollen und geschlechtstypisches Verhalten*. Stuttgart: Kohlhammer.
Allingham, M., & Sandmo, A. (1972). Income tax evasion: A theoretical analysis. *Journal of Public Economics, 1*, 323–338.
Anselmi, D. L., & Law, A. L. (Eds.). (1998). *Questions of gender: Perspectives and paradoxes*. Boston, MA: McGraw-Hill.
Apicella, C. L., Dreber, A., Campbell, B., Gray, P. B., Hoffman, M., & Little, A. C. (2008). Testosterone and financial risk preferences. *Evolution and Human Behavior, 29*, 384–390.
Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173–1182.
Bazart, C., & Pickhardt, M. (2009). *Fighting income tax evasion with positive rewards: Experimental evidence*. LAMETA, University of Montpellier Working Papers, 09-01. Available from: <http://ideas.repec.org/p/lam/wpaper/09-01.html>.
Bem, S. (1981). *Bem sex-role inventory*. Palo Alto, CA: Consulting Psychologists Press.
Betz, M., O'Connell, L., & Shepard, J. M. (1989). Gender differences in proclivity for unethical behavior. *Journal of Business Ethics, 8*, 321–324.
Bonnes-Dobrowolny, M., & Vicarelli, G. (1982). Costruzione di un personal attributes questionnaire italiano per la misura dell'androginia psicologica. *Communicazioni Scientifiche di Psicologia Generale, 9*, 105–144.
Burnham, T. C. (2007). High-testosterone men reject low ultimatum game offers. *Proceedings of the Royal Society B, 274*, 2327–2330.
Byrnes, J. P., Miller, D. C., & Schafer, W. D. (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin, 125*, 367–383.
Chung, J., & Trivedi, V. U. (2003). The effect of friendly persuasion and gender on tax compliance behavior. *Journal of Business Ethics, 47*, 133–145.
Coates, J. M. (2009). Reply to millet: Digit ratios and high frequency trading. *Proceedings of the National Academy of Sciences of the USA, 106*, E31.

Coates, J. M., Gurnell, M., & Rustichini, A. (2009). Second-to-fourth digit ratio predicts success among high-frequency financial traders. *Proceedings of the National Academy of Sciences of the USA, 106*, 623–628.
Cohen-Bendahan, C. C. C., van de Beek, C., & Berenbaum, S. A. (2005). Prenatal sex hormone effects on child and adult sex-typed behavior: Methods and findings. *Neuroscience and Biobehavioral Reviews, 29*, 353–384.
Croson, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature, 47*, 448–474.
Csathó, Á., Osváth, A., Bicsák, É., Karádi, K., Manning, J. T., & Kállai, J. (2003). Sex role identity related to the ratio of second to fourth digit length in women. *Biological Psychology, 62*, 147–156.
Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. Hillsdale, NJ: Lawrence Erlbaum.
Friedland, N., Maital, S., & Rutenberg, A. (1978). A simulation study of income tax evasion. *Journal of Public Economics, 10*, 107–116.
Gentile, D. A. (1998). Just what are sex and gender, anyway? A call for a new terminological standard. In D. L. Anselmi & A. L. Law (Eds.), *Questions of gender: Perspectives and paradoxes* (pp. 14–17). Boston, MA: McGraw-Hill.
Gerxhani, K. (2007). Explaining gender differences in tax evasion: The case of Tirana, Albania. *Feminist Economics, 13*, 119–155.
Grasso, L., & Kaplan, S. E. (1998). An examination of ethical standards. *Journal of Accounting Education, 16*, 85–100.
Hasseldine, J. (1999). Gender differences in tax compliance. *Asia-Pacific Journal of Taxation, 3*, 73–89.
Hines, M. (2004). Androgen, estrogen, and gender: Contribution of the early hormone environment to gender-related behavior. In A. H. Eagly, A. E. Beall, & R. J. Sternberg (Eds.), *The psychology of gender* (2nd ed., pp. 9–37). New York: Guilford Press.
Jackson, B. R., & Milliron, V. C. (1986). Tax compliance research: Findings, problems, and prospects. *Journal of Accounting Literature, 5*, 125–165.
Jianakoplos, N. A., & Bernasek, A. (1998). Are women more risk averse? *Economic Inquiry, 36*, 620–630.
Kastlunger, B., Kirchler, E., Mittone, L., & Pitters, J. (2009). Sequences of audits, tax compliance, and taxpaying strategies. *Journal of Economic Psychology, 30*, 405–418.
Kinsey, K. A. (1992). Deterrence and alienation effects of IRS enforcement: An analysis of survey data. In J. Slemrod (Ed.), *Why people pay taxes: Tax compliance and enforcement* (pp. 259–285). Ann Arbor, MI: University of Michigan Press.
Kirchler, E. (2007). *The economic psychology of tax behaviour*. Cambridge, MA: Cambridge University Press.
Kirchler, E., & Maciejovsky, B. (2001). Tax compliance within the context of gain and loss situations, expected and current asset position, and profession. *Journal of Economic Psychology, 22*, 173–194.
Lewis, A., Carrera, S., Cullis, J., & Jones, P. (2009). Individual, cognitive and cultural differences in tax compliance. UK and Italy compared. *Journal of Economic Psychology, 30*, 431–445.
Lippa, R. A. (1991). Some psychometric characteristics of gender diagnosticity measures: Reliability, validity, consistency across domains, and relationship to the big five. *Journal of Personality and Social Psychology, 61*, 1000–1011.
Lippa, R. A. (1995). Gender-related individual differences and psychological adjustment in terms of the big five and circumplex models. *Journal of Personality and Social Psychology, 69*, 1184–1202.
Lippa, R. A. (2002). Gender-related traits of heterosexual and homosexual men and women. *Archives of Sexual Behavior, 31*, 83–98.
Lippa, R. A. (2006). Finger lengths, 2D:4D ratios and their relation to gender-related personality traits and the big five. *Biological Psychology, 71*, 116–121.
Lippa, R. A., & Connelly, S. (1990). Gender diagnosticity: A new Bayesian approach to gender-related individual differences. *Journal of Personality and Social Psychology, 59*, 1051–1065.
McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1*, 30–46.
Manning, J. T. (2002). *Digit ratio: A pointer to fertility, behaviour, and health*. New Brunswick, NJ: Rutgers University Press.
Mason, R., & Calvin, L. D. (1978). A study of admitted income tax evasion. *Law and Society Review, 13*, 73–89.
Mazur, A., & Booth, A. (1998). Testosterone and dominance in men. *Behavioral and Brain Sciences, 21*, 353–397.
Meier-Pesti, K., & Penz, E. (2008). Sex or gender? Expanding the sex-based view by introducing masculinity and femininity as predictors of financial risk taking. *Journal of Economic Psychology, 29*, 180–196.
Millet, K. (2009). Low second-to-fourth-digit ratio might predict success among high-frequency financial traders because of a higher need for achievement. *Proceedings of the National Academy of Sciences of the USA, 106*, E30.
Millet, K., & Dewitte, S. (2006). Second to fourth digit ratio and cooperative behavior. *Biological Psychology, 71*, 111–115.
Millet, K., & Dewitte, S. (2008). A subordinate status position increases the present value of financial resources for low 2D:4D men. *American Journal of Human Biology, 20*, 110–115.
Millet, K., & Dewitte, S. (2009). The presence of aggression cues inverts the relation between digit ratio (2D:4D) and prosocial behaviour in a dictator game. *British Journal of Psychology, 100*, 151–162.
Peters, M., Tan, Ü., Kang, Y., Teixeira, L., & Mandal, M. (2002). Sex-specific finger-length patterns linked to behavioral variables: Consistency across various human populations. *Perceptual and Motor Skills, 94*, 171–181.
Porcano, T. M. (1988). Correlates of tax evasion. *Journal of Economic Psychology, 9*, 47–67.
Rammsayer, T. H., & Troche, S. J. (2007). Sexual dimorphism in second-to-fourth digit ratio and its relation to gender-role orientation in males and females. *Personality and Individual Differences, 42*, 911–920.
Richards, P., & Tittle, C. R. (1981). Gender and perceived chances of arrest. *Social Forces, 59*, 1182–1199.
Ronay, R., & von Hippel, W. (2009). Power, testosterone, and risk-taking. *Journal of Behavioral Decision Making*. doi:10.1002/bdm.671.
Sapienza, P., Zingales, L., & Maestripieri, D. (2009). Gender differences in financial risk aversion and career choices are affected by testosterone. *Proceedings of the National Academy of Sciences of the USA, 106*, 15268–15273.
Schmukle, S. C., Liesenfeld, S., Back, M. D., & Egloff, B. (2007). Second to fourth digit ratios and the implicit gender self-concept. *Personality and Individual Differences, 43*, 1267–1277.
Smith, K. W. (1992). Reciprocity and fairness: Positive incentives for tax compliance. In J. Slemrod (Ed.), *Why people pay taxes: Tax compliance and enforcement* (pp. 223–250). Ann Arbor, MI: University of Michigan Press.
Spicer, M. W., & Becker, L. A. (1980). Fiscal inequity and tax evasion: An experimental approach. *National Tax Journal, 33*, 171–175.
Spicer, M. W., & Hero, R. E. (1985). Tax evasion and heuristics: A research note. *Journal of Public Economics, 26*, 263–267.
Storms, M. D. (1979). Sex role identity and its relationships to sex role attributes and sex role stereotypes. *Journal of Personality and Social Psychology, 37*, 1779–1789.
Sunden, A. E., & Surette, B. J. (1998). Gender differences in the allocation of assets in retirement savings plans. *American Economic Review, 88*, 207–211.
Torgler, B., & Schneider, F. (2004). *Does culture influence tax morale? Evidence from different European countries*. CREMA Working papers, 2004-17. Available from: <http://www.crema-research.ch/papers/2004-17.pdf>.
Torgler, B., & Valev, N. T. (2006). *Women and illegal activities: Gender differences and women's willingness to comply over time*. CREMA Working papers, 2006-15. Available from: <http://www.crema-research.ch/papers/2006-15.pdf>.
Unger, R. K. (1979). Toward a redefinition of sex and gender. *American Psychologist, 34*, 1085–1094.
Unger, R. K., & Crawford, M. (1998). Sex and gender: The troubled relationship between terms and concepts. In D. L. Anselmi & A. L. Law (Eds.), *Questions of gender: Perspectives and paradoxes* (pp. 18–21). Boston, MA: McGraw-Hill.
van den Bergh, B., & Dewitte, S. (2006). Digit ratio (2D:4D) moderates the impact of sexual cues on men's decisions in ultimatum games. *Proceedings of the Royal Society B, 273*, 2091–2095.
Vogel, J. (1974). Taxation and public opinion in Sweden: An interpretation of recent survey data. *National Tax Journal, 27*, 499–513.
Voracek, M., & Dressler, S. G. (2006). Lack of correlation between digit ratio (2D:4D) and Baron–Cohen's "Reading the Mind in the Eyes" test, empathy, systemising, and autism-spectrum quotients in a general population sample. *Personality and Individual Differences, 41*, 1481–1491.
Voracek, M., & Loibl, L. M. (2009). Scientometric analysis and bibliography of digit ratio (2D:4D) research, 1998–2008. *Psychological Reports, 104*, 922–956.

Voracek, M., Manning, J. T., & Dressler, S. G. (2007). Repeatability and interobserver error of digit ratio (2D:4D) measurements made by experts. *American Journal of Human Biology, 19*, 142–146.

Wenzel, M. (2002). The impact of outcome orientation and justice concerns on tax compliance. The role of taxpayers' identity. *Journal of Applied Psychology, 87*, 629–645.

Wilson, J., & Sheffrin, S. (2005). Understanding surveys of taxpayer honesty. *Finanzarchiv, 61*, 256–274.

Please cite this article in press as: Kastlunger, B., et al. Sex differences in tax compliance: Differentiating between demographic sex, gender-role orientation, and prenatal masculinization (2D:4D). *Journal of Economic Psychology* (2010), doi:10.1016/j.joep.2010.03.015