Fetal Testosterone (2D:4D) as Predictor of Cognitive Reflection

Antoni Bosch-Domènech
Pablo Brañas-Garza
Antonio M. Espín

May 2013

Barcelona GSE Working Paper Series

Working Paper nº 698

Fetal testosterone (2D:4D) as predictor of cognitive reflection*

Antoni Bosch-Domènech[a], Pablo Brañas-Garza[b,1], Antonio M. Espín[c]
[a]Department of Economics and Business, Universitat Pompeu Fabra and Barcelona Graduate School of Economics, 08005 Barcelona, Spain; [b]Middlesex University London, Business School, Hendon Campus, The Burroughs, London NW4 4BT ; [c] Facultad de Ciencias Económicas y Empresariales, Campus de la Cartuja s/n, 18071 Granada, Spain.

Corresponding author: Antoni Bosch-Domènech, Department of Economics and Business, Universitat Pompeu Fabra, Ramon Trias Fargas 25, 08005 Barcelona, Spain.

Classification: Economic Science, Psychological and Cognitive Sciences

* The authors' names are placed in alphabetic order. All of them contributed equally to the paper.

Abstract

The Cognitive Reflection Test (CRT) is a test introduced by S. Frederick (2005) Cognitive reflection and decision making, *J Econ Perspect* 19(4): 25-42. The task is designed to measure the tendency to override an intuitive response that is incorrect and to engage in further reflection that leads to the correct response. The consistent sex differences in CRT performance may suggest a role for gonadal hormones, particularly testosterone. A now widely studied putative marker for fetal testosterone is the second-to-fourth digit ratio (2D:4D). This paper tests to what extent 2D:4D, as a proxy for prenatal exposure to testosterone, can predict CRT scores in a sample of 623 students. After controlling for sex, we observe that a lower 2D:4D (reflecting a higher exposure to testosterone) is significantly associated with a higher number of correct answers. The result holds for both hands' 2D:4Ds. In addition, the effect appears to be sharper for females than for males. We also control for patience and math proficiency, which are significantly related to performance in the CRT. But the effect of 2D:4D on performance in CRT is not reduced with these controls, implying that these variables are not mediating the relationship between digit ratio and CRT.

Keywords: Cognitive Refection Test, 2D:4D, fetal testosterone, patience, sex.

Introduction

The Cognitive Reflection Test (CRT) is a three-item test introduced by Frederick (1). The task, of an algebraic nature, is designed to measure the tendency to override an intuitive response that is incorrect and to engage in further reflection that leads to the correct response. When answering the test, many people give the first response that comes to mind without thinking further and not realizing that it cannot be the right answer.* Cognitive ability is no guarantee against making the error. What makes the CRT different from problem-solving or math tests is that the latter tests do not usually trigger a plausible intuitive response that must be overridden; instead, the respondent is lost in thought trying to solve the problem correctly. As Kahneman and Frederick (2) made clear, the framework of an incorrectly primed initial response that must be overridden fits in nicely with currently popular dual-process frameworks, one emotional/impatient and the second one deliberative/patient.† See refs. 4-8.

Frederick (1) observed that with as few as three items his CRT was able to predict performance on measures of temporal discounting, risk preference, and the tendency to choose high-expected-value gambles. A large literature has developed about the relation between CRT and performance, but the data have proved to be inconsistent in some instances. See refs. 9-13.

Yet, the better responses to the CRT by males appear to be a robust result (1, 13). The sex differences in CRT answers may suggest a role for gonadal hormones, particularly testosterone.‡ A now widely studied putative marker for fetal testosterone is the second-to-fourth digit ratio (2D:4D). See refs.

* One item from the CRT is: "*A bat and a ball cost $1.10. The bat costs $1.00 more than the ball. How much does the ball cost? _____ Cents.*" A glib, incorrect, and frequent answer is 10 cents; the correct answer is 5 cents.

† The dual process of emotional/deliberative mental systems has received different names: Fast and slow thinking, hot and cold, locomotion and assessment, automatic and controlled thought (3).

‡ Traits that may be linked with prenatal exposure to testosterone expression are, among others, spatial/mathematical skills (14, 15); performance in computer science (16); heightened attention to detail, intensified focus, and narrow interests (17); less emotion recognition, eye contact and social sensitivity, a poorer ability to judge what others are thinking or feeling, lack of empathy (18).

19-21. Earlier studies that have stood the test of replication have reported that 2D:4D varies by sex and ethnicity but that male 2D:4D tends to be smaller than female 2D:4D in all ethnic groups and the effect is strongest in the right hand (22). These differences emerge prenatally and appear to be stable during the developing years (22-25). Digit ratio has been associated with cognitive abilities (26); impulsivity (27); aggression (28-30) and risk-taking (26 and 31-34), among other effects on personality and cognition. While the conjecture that the sex difference in human digit ratios reflects prenatal androgen exposure has been widely confirmed (35), probably the digit ratios imperfectly reflect prenatal testosterone stimulation[§], so one must gather a sufficiently large sample to relate that stimulation to human behavior statistically.

The purpose of the paper is to test to what extent 2D:4D, as a proxy for prenatal exposure to testosterone, correlates with the CRT results in a non-random sample of 623 students.

Results and discussion

1 - 2D:4D

Descriptive statistics of the 2D:4D measurements, including tests of normality, are presented in Table 1. The results are displayed separately for males and females and for left and right hands. We find no significant departure from normality of the D2:D4 data except in the case of males' right hand, for which the normality test reaches marginal significance (p = 0.099) due to a non-normally skewed distribution (p = 0.034).

The digit ratio is significantly higher in the left hand than in the right hand for both men ($t(259)$ = 3.2708, p = 0.001; two-sided t-test) and women ($t(362)$= 2.4716, p = 0.014). In line with previous literature (37-39), the digit ratio was found to be lower for men than for women (right hand: $t(621)$ = 4.4661, p = 0.000, left hand: $t(621)$ = 3.8079, p = 0.000).

Table 1. Descriptive statistics of 2D:4D

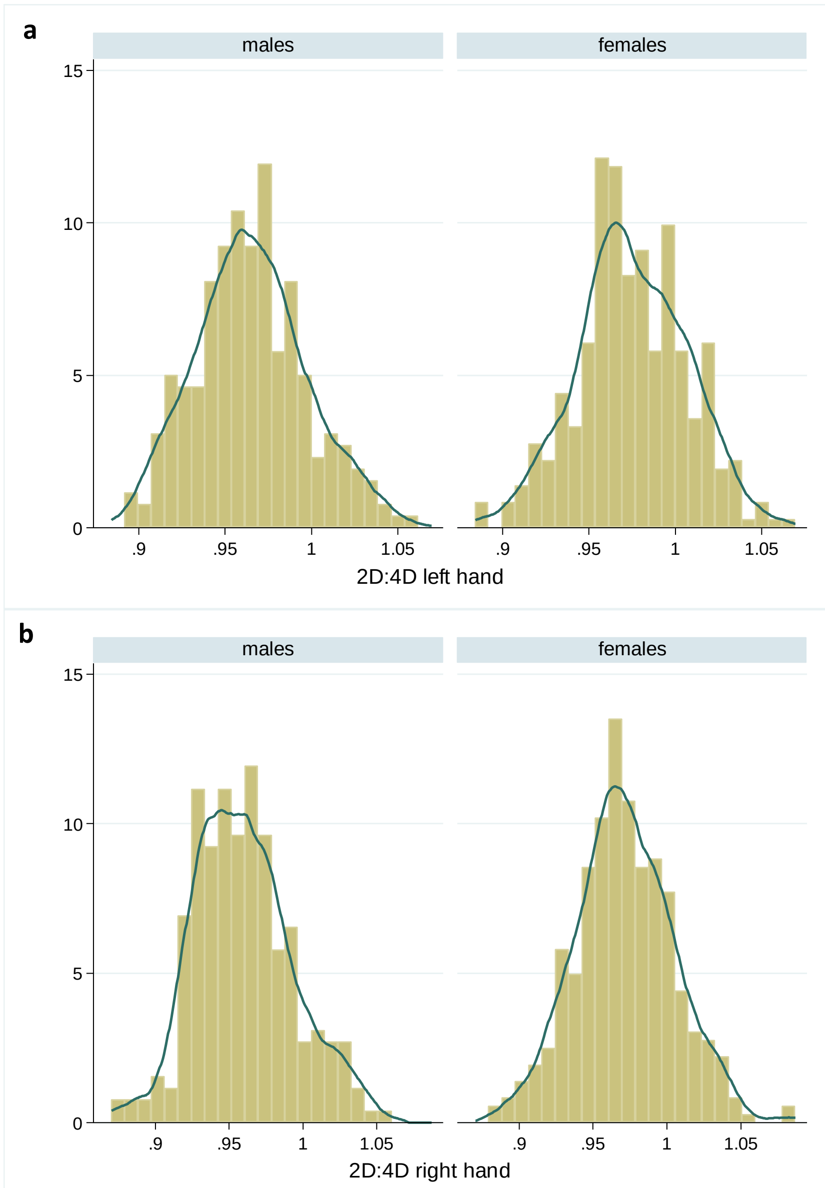
	males		females	
	left	right	left	right
mean	0.9651	0.9597	0.9749	0.9717
sd	0.0317	0.0333	0.0316	0.0332
sem	0.0020	0.0021	0.0017	0.0017
median	0.9639	0.9585	0.9737	0.9695
skewness	0.2403	0.321	-0.013	0.180
p-value	0.109	0.034	0.915	0.156
kurtosis	2.809	3.026	2.932	3.181
p-value	0.617	0.763	0.922	0.394
normal (*Chi*^{2})	2.84	4.63	0.02	2.75

§ Prenatal estradiol seems also to have an effect on digit length (36).

p-value	0.241	0.099	0.989	0.253

Figure 1 reports the histogram and kernel density estimation of 2D:4D in our sample. The results are displayed separately for males and females and for the left hand (panel *a*) and right hand (panel *b*).

Figure 1. Distribution of 2D:4D: Histogram and kernel density



2- 2D:4D and CRT

The results of the CRT appear in Table 2. The upper part of the table reports, for each question, the percentage of males and females who answered it correctly and the significance level of the difference between sexes (two-sided Fisher's exact test). Men were significantly more likely than women to answer correctly each of the three questions (although for question 1 the difference is only marginally significant). The mean (±sem) number of correct responses in the CRT was 0.958 ± 0.064 for males and 0.584 ± 0.045 for females (Cohen's *d* = 0.3941).

Table 2. CRT: % of correct answers by sex

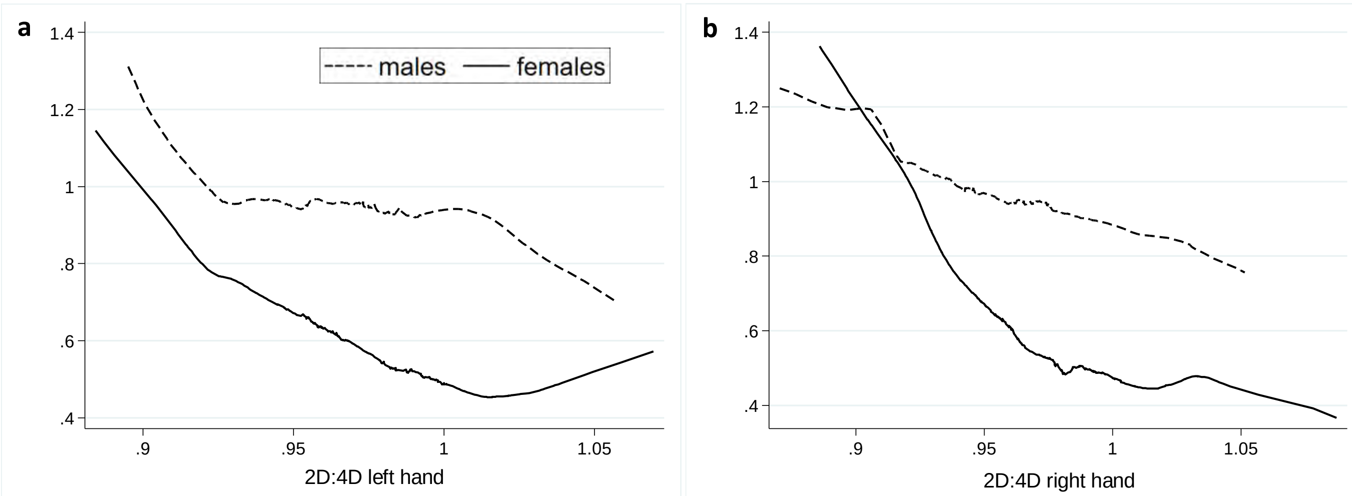
	Males (%)	Females (%)	*p-value*
CRT-item 1	35.77	29.20	0.098
CRT-item 2	25.77	10.47	0.000
CRT-item 3	34.23	18.73	0.000
0 correct answers	43.46	61.43	
1 correct answer	28.85	23.97	
2 correct answers	16.15	9.37	
3 correct answers	11.54	5.23	

p-values from two-sided Fisher's exact tests for the difference in proportions.

The bottom part of the table reports the distribution of the number of correct answers for males and females: 27.69% of males had two or three correct answers in the CRT, while this percentage shrinks to 14.60% for females, and 11.54% of males and 5.23% of females answered correctly all the three CRT questions. A notable fraction of the subject pool (43.46% of males and 61.43% of females) was unable to solve any of the referred questions.

The relationship between the subjects' performance in the CRT and their 2D:4D is shown in Figure 2. Smoothed curves were fit using locally weighted regressions (LOWESS smoothing) with a standard, conservative bandwidth of 0.8. For both sexes, we observe a *negative relationship* between the number of correct answers in the CRT and both the left-hand (panel *a*) and the right-hand (panel *b*) 2D:4D. In addition, the effect of the DR on the number of correct answers in the CRT appears to be sharper for females than for males.

Figure 2. LOWESS smoothing: Cognitive reflection as a function of 2D:4D



Column (1) of Table 3 presents estimates of an ordered probit regression for the effects of 2D:4D and sex on the number of correct answers to the CRT (left panels refer to the left hand and right panels to the right hand). Zero-order correlations between all the variables used are reported, separately for males and females, in Table A1 in the Appendix.

Table 3. The impact of 2D:4D on CRT

	a) Left hand				*b) Right hand*			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
2D:4D	-3.225**	-1.550	-3.829***	-1.869	-4.572***	-1.827	-4.977***	-2.431
	(1.465)	(2.174)	(1.483)	(2.177)	(1.410)	(2.076)	(1.424)	(2.082)
Female	-0.424***	2.543	-0.336***	3.186	-0.407***	4.499*	-0.321***	4.276
	(0.094)	(2.850)	(0.096)	(2.870)	(0.094)	(2.733)	(0.096)	(2.751)
2D:4D x Female		-3.062		-3.635		-5.090*		-4.771*
		(2.940)		(2.961)		(2.834)		(2.854)
Math			0.265***	0.268***			0.265***	0.265***
			(0.061)	(0.061)			(0.061)	(0.061)
Impatience			-0.041**	-0.041**			-0.041**	-0.039**
			(0.019)	(0.019)			(0.019)	(0.019)
log likelihood	-695.863	-695.321	-683.461	-682.707	-692.993	-691.377	-680.641	-679.241
Chi^2	28.57***	29.65***	53.37***	54.88***	34.31***	37.54***	59.01***	61.81***
pseudo R^2	0.0201	0.0209	0.0376	0.0386	0.0242	0.0264	0.0415	0.0435
N	623	623	623	623	623	623	623	623

Note: Ordered probit estimates. Standard errors in brackets. *,**,*** indicate significance at the 0.1, 0.05 and 0.01 levels, respectively.

A lower 2D:4D is significantly associated with a higher number of correct answers (left hand: p = 0.028; right hand: p = 0.001), and males had significantly more correct answers than females (p = 0.000). Interaction effects are shown in column (2). There is a marginally significant interaction between right-hand 2D:4D and sex (p = 0.072), indicating that the negative impact of 2D:4D on CRT is more pronounced for females. Wald tests on the coefficients of that model indicate that the effect is significant for females ($Chi^2(1) = 12.82$, p = 0.000) but not for males ($Chi^2(1) = 0.77$, $p > 0.3$). No significant interaction effect is found for the left-hand 2D:4D ($p > 0.2$), although the sign of the interaction term is the same as for the right hand (i.e., more pronounced effect for females).

As mentioned above, the negative impact of 2D:4D on CRT is more pronounced for females than for males. A differential impact of 2D:4D between sexes has often been reported in the literature.[**] Frederick (1) also observes that CRT scores are more highly correlated with time preferences for women than for men. This may suggest that some of the effect of 2D:4D on the CRT is due to time preference or impatience. After all, according to a dual-process approach, answering correctly the CRT appears to require that the deliberative/patient mind overrules the intuitive/impatient response. Similarly one could posit that some of the effect of 2D:4D on the CRT may signal mathematical ability, since the CRT questions, although simple, have an algebraic content. To disentangle whether the effect of 2D:4D on CRT is in fact capturing the impact of *mathematical ability* or a *degree of impatience*, we extend our analysis to account for these two factors.

We now estimate the effects of 2D:4D and sex, as before, but controlling for the effect of math proficiency, as measured by the number of correct answers to the mathematical test, and for the effect of impatience, as measured by the number of impatient answers in the time preference task. The results appear in columns (3) and (4) of Table 3 (for both the left and right hands).

As in Frederick (1), we find that impatience is negatively and significantly related to performance in the CRT (*ps* < 0.05). As expected, mathematical ability is a positive and strong determinant of CRT scores (*ps* < 0.01). Yet, there is an interesting insight obtained from these regressions: The effect of 2D:4D on CRT *is not reduced* (it even increases slightly; right hand: p = 0.000, left hand: p = 0.010; column (3)) when controlling for the performance in the math and impatience tests. This implies that these variables are not mediating the relationship between digit ratio and CRT. It appears, therefore, that the effect of 2D:4D captures a component of the determinants of the subjects' performance in the CRT that is *different* from the effect of sex, performance in a simple mathematical test, and impatience.[††] So further research should try to test whether other factors, like enhanced persistence in

[**] For a significant differential impact, see, e.g., on visual-spatial abilities (40,41); on musical abilities (42); on numerical ability/literacy (43,44)); on sensation seeking (45,46).

[††] Some may argue that being more reflective, as measured by the CRT, leads to less impatient behavior in the time preferences task, rather than the opposite causal way. To alleviate this concern, we performed partial correlations between CRT scores and each of the explanatory variables, while keeping the other variables constant: the significance levels remain nearly identical to those reported in Table 3 (available upon request from the authors). And, clearly, the causality of the main relationship (that is, prenatal testosterone exposure impacts on CRT scores) cannot be reversed.

an effort, or increased ability not to be distracted by irrelevant information, or higher "need for achievement" (47), may mediate the effect of fetal testosterone on CRT. [‡‡]

Methods

In October 2011, 927 first-year students at the College of Business and Economics of the University of Granada (Spain) were asked to participate in a survey-experiment at the Laboratory of Experimental Economics, EGEO. Participation was voluntary and the number of participants ended up being 659 (71% of the population), distributed in 27 sessions. All subjects gave written informed consent to participate. We excluded from the sample those observations with missing values in any of the variables used in this paper. To ensure ethnic homogeneity, three non-Caucasian subjects were also excluded from the sample. The resulting sample was composed of 623 ethnically homogeneous subjects (260 males; mean age 19.1±2.3 (sd)). During a session, using a computer-based system, participants were asked to complete several questionnaires on their socio-demographic characteristics, were tested for their time-discounting attitudes, and answered a math test with four questions, three of which are straightforward. After responding to the computer-based questionnaires, participants answered the CRT's three questions using paper and pencil. No time pressure was imposed in any of the processes.

To test the participants for their time-discounting attitudes, they were presented with two series of intertemporal decisions involving hypothetical monetary rewards.[§§] Participants faced a total of six decisions in each subtask. In the first decision of the first subtask, participants had to choose between €5 to be received "today" (sooner option) and €5 to be received "tomorrow" (later option). The remaining five decisions kept the sooner reward constant while increasing the later reward, in this order: €6, €7, €8, €9, €10. In the first decision of the second subtask, participants had to choose between €150 to be received in one month time (sooner option) and €150 to be received in seven months' time (later option). Again, keeping the sooner option constant, the later options in the remaining five decisions were: €170, €190, €210, €230, €250.[***] The total number of impatient choices (from 0 to 12) is our measure of impatience.[†††]

[‡‡] In our sample, DR does not correlate significantly with the number of correct answers in the math test (*ps*>0.2; see Table A1), except in the case of females' left hand (p = 0.034). That the latter relationship is positive may explain why the negative impact of DR on the CRT score is even stronger when controlling in the regressions of Table 3 for the number of correct answers in the math test. It is arguable that the simplicity of the latter or the different procedures used (the math test was embedded in a long questionnaire while the CRT was presented as a separate task) have influenced the results. Indeed, it has been hypothesized that higher prenatal exposure to testosterone might predict a higher "need for achievement" (47), which would possibly be more prominent in more self-motivating, complicated or salient tasks.

[§§] Previous studies have shown that the distribution of individual choices in time preference tests is not significantly altered by the existence of real (*vs*. hypothetical) incentives, neither within nor between subjects (50-53).

[***] For experiments on time preference using more extended tasks with similar delays and rewards, see refs. 53-55.

[†††] We excluded from the sample the 13 subjects making inconsistent choices in any of the subtasks (i.e., non-monotonic patterns or multiple switching from sooner to later reward).

The questions for the CRT and the math test are presented in the Appendix. We describe below the results of these two tests by the number of correct answers to them.[‡‡‡]

After taking the tests, the participants were asked one by one to have their two hands scanned using a high-resolution scanner (Canon Slide 90) and their fingers were measured from the middle of the basal crease to the tip of the finger using Photoshop. Computer-assisted measurements of 2D:4D from scanned pictures have been found to be more precise and reliable than measurements using other methods (57-58). The 2D:4D of the scanned pictures was measured twice for each hand at an interval of one month by the same experienced measurer (not involved in this paper). These measurements displayed a high repeatability (right hand: intraclass correlation coefficient (ICC)= 0.9566, p = 0.000, left hand: ICC = 0.9440, p = 0.000) and were averaged to obtain a single value of the 2D:4D ratio for each hand.

Conclusion

From the results presented above we have to conclude that fetal hormone exposure, expressed in its putative marker 2D:4D, has a significant and positive effect on how females and, to a more ambiguous degree, males answer the CRT. In plain words, we observe an association between DR and CRT scores, which suggests a relation between higher levels of fetal testosterone and attention, concentration, diligence or whatever traits that, beyond competence in algebra, facilitate the overriding of the intuitive but incorrect responses to the test. In our large sample of first-year college students some do think through the intuitive answer while others do not. 2D:4D can help to predict who will and who will not, especially among women. Our results show that women with less prenatal exposure to testosterone (to the extent that it is expressed in 2D:4D) do poorly compared with women with more prenatal exposure to testosterone.

It is known that early androgen surges exert an organizational influence on brain development which indicates that fetal testosterone in humans can act as a programming mechanism that influences behavior later in life.[§§§] Admittedly, trying to pin down differences in the CRT answers to one single factor, fetal testosterone, is simplistic and may eventually lead to conflicting, erratic or inconclusive results (indeed, from the pseudo-R^2 values reported in Table 3, it can be observed that much of the variation remains unexplained in our regressions). While 2D:4D is a fixed and predetermined variable, other processes influencing behavior may have occurred or may even be occurring while subjects take the test. Coates (49) conjectures a "preparation for the test effect" and a "winner effect" (that in our test may result from the satisfaction of answering correctly the first question in the CRT) resulting in a variation in circulating hormones that may distort the predictive power of the 2D:4D biometric

[‡‡‡] The math questions come from Section K of *Encuesta de Protección Social*, 2009 by the Government of Chile. Participants were also asked to play some economic games. For the details of the survey-experiment, with another sample, see refs. 56.

[§§§] Lombardo *et al*. (48), found that preadolescent boys who, as fetuses, had unusually high levels of testosterone in the amniotic fluid that surrounded them had more gray matter in a site in the right hemisphere that contributes to the solution of spatial reasoning problems, but less gray matter in sites in the left hemisphere that contribute to language.

measurements. Still, the robust effect of both hands' 2D:4D ratios on subjects' answers to the CRT, which is not mediated by their answers to the impatience or basic math tests, should encourage further controlled experiments to pin down why individuals exposed to a larger than average amount of testosterone in utero offer better, more reasoned, solutions in the CRT twenty years after the fact.

Ethics statement. All participants in the experiments reported in the manuscript were informed about the content of the experiment before they participated and provided written consent. Besides, their anonymity was always preserved (in agreement with the Spanish Law 15/1999 for Personal Data Protection) by assigning them randomly a numerical code, which would identify them in the system. No association was ever made between their real names and the results. As it is standard in socio-economic experiments, no ethic concerns are involved other than preserving the anonymity of participants. This procedure was checked and approved by the Vice dean of Research of the School of Economics of the University of Granada, the institution hosting the experiment.

Acknowledgements. This paper has been benefited by comments and suggestions of Mark Coulson, Tom Dickins, Praveen Kujal, Kobe Millet and Martin Voracek. The authors thank Levent Neyse for his help in data collection. The first author acknowledges financial aid from Spanish Ministerio de Educación y Ciencia (ECO2011-2529) and from the Spanish Ministerio de Economía through the Severo Ochoa Programme for Centres of Excellence in R&D (SEV-2011-0075). The second and the third authors acknowledge financial support from Ministerio Educación y Ciencia (ECO2010-17049), Fundación Ramón Areces R+D 2011 and Junta de Andalucía-Excelencia (P07.SEJ.02547).

References

1. Frederick S (2005) Cognitive reflection and decision making. *J Econ Perspect* 19(4): 25-42.
2. Kahneman D, Frederick S (2002) in *Heuristics and Biases: The Psychology of Intuitive Judgment,* eds Gilovich T, Griffin D, Kahneman D (Cambridge University Press, New York), pp 49-81.
3. Camerer C, Loewenstein G, Prelec D (2005) Neuroeconomics: how neuroscience can inform economics. *J Econ Lit* 43(1): 9-64.
4. Alter AL, Oppenheimer DM, Epley N, Eyre RN (2007) Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *J Exp Psychol Gen* 136(4): 569-576.
5. Bernheim BD, Antonio R (2004) Addiction and cue-triggered decision processes. *Am Econ Rev* 94(5):1558-1590.
6. Brocas I, Juan C (2008) The brain as a hierarchical organization. *Am Econ Rev* 98(4): 1312-1346.
7. Fudenberg D, David KL (2006) A dual-self model of impulse control. *Am Econ Rev* 96(5):1449-1476.
8. Loewenstein G, Ted OD (2005). Animal spirits: Affective and deliberative processes in economic behavior. Cornell University Mimeograph.
9. Cokely ET, Kelley, CM (2009) Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation. *Judgm Decis Mak* 4(1): 20-33.
10. Campitelli G, Labollita M (2010) Correlations of cognitive reflection with judgments and choices. *Judgm Decis Mak* 5(3): 182-191.
11. Koehler DJ, James G (2010) Probability matching and strategy availability. *Mem & Cogn* 38(6): 667-676.
12. Toplak ME, Richard FW, Keith ES (2011) The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Mem& Cogn* 39(7): 1275-1289.
13. Oechssler J, Andreas R, Patrick WS (2009) Cognitive abilities and behavioral biases. *J Econ Behav*

Organ 72(1): 147-152.

14. Geschwind N, Galaburda AM (1985) Cerebral lateralization – Biological mechanisms, associations and pathology: A hypothesis and a program for research. *Arch Neurol* 42(5): 428–459
15. Grimshaw GM (1995) Relations between prenatal testosterone and cerebral lateralization in children. *Neuropsy* 9(1): 68–79.
16. Brosnan M, Gallop V, Iftikhar N, Keogh E (2011) Digit ratio (2D: 4D), academic performance in computer science and computer-related anxiety. *Pers Individ Dif 51*(4): 371-375.
17. Baron-Cohen S, Knickmeyer R, Belmonte M (2005) Sex differences in the brain: implications for explaining autism. *Science* 310(5749): 819-823.
18. Baron-Cohen S, Lutchmaya S, Knickmeyer R (2004) *Prenatal testosterone in mind: Studies of amniotic fluid* (MIT Press, Cambridge, MA).
19. Voracek M, Loibl LM (2009) Scientometric analysis and bibliography of digit ratio (2D:4D) research, 1998–2008. *Psychol Rep* 104(3): 922–956.
20. Fisher, Helen E, Jonathan R, Heide DI, Daniel M (2010) The second to fourth digit ratio: A measure of two hormonally-based temperament dimensions. *Pers Individ Dif* 49(7): 773-777.
21. Zheng Z, Martin J (2011) Developmental basis of sexually dimorphic digit ratios. *Proc Natl Acad Sci USA* 108(39): 16289-16294.
22. Manning JT (2002) *Digit ratio: a pointer to fertility, behavior, and health*, Rutgers University Press, New Jersey.
23. Voracek M, Manning JT, Dressler SG (2007) Repeatability and Interobserver Error of Digit Ratio (2D:4D) Measurements Made by Experts. *Am J Hum Bio* 19(1):142–146.
24. Manning JT, Churchill AJG, Peters M (2007) The effects of sex, ethnicity, and sexual orientation on self-measured digit ratio (2D:4D). *Arch Sex Behav* 36(2): 223–233.
25. Manning JT, Fink B (2008) Digit Ratio (2D:4D), Dominance, Reproductive Success, Asymmetry, and Sociosexuality and in the BBC Internet Study. *Am J Hum Bio* 20(4): 451-461.
26. Brañas-Garza P, Teresa GM, Roberto HG (2012) Cognitive effort in the Beauty Contest Game. *J Econ Behav Organ* 83(2): 254-260.
27. Hanoch Y, Gummerum M, Rolison J (2012). Second-to-Fourth Digit Ratio and Impulsivity: A Comparison between Offenders and Non-offenders. *PloS one, 7*(10), e47140. doi:10.1371/journal.pone.0047140.
28. Hampson E, Ellis CL, Tenk CM (2008) On the relation between 2D:4D and sex-dimorphic personality traits. *Arch Sex Behav* 37(1): 133-144.
29. Coyne SM, Manning JT, Ringer L, Bailey L (2007) Directional Asymmetry in digit ratio (2D:4D) predict indirect aggression in women. *Pers Individ Dif* 43(4): 865-872.
30. Bailey AA, Hurd PL (2005) Finger length ratio (2D:4D) correlates with physical aggression in men but not in women. *Biol Psychol* 68(3):215-222.
31. Coates JM, Gurnell M, Rustichini (2009) A. Second-and fourth digit ratio predicts success among higher frequency financial traders. *Proc Natl Acad Sci USA* 106(2): 623-628.
32. Stenstrom E, Saad G, Nepomuceno MV, Mendenhall Z (2011) Testosterone and domain-specific risk: Digit ratios (2D:4D and rel2) as predictors of recreational, financial, and social risk-taking behaviors. *Pers Individ Dif* 51(4):412-416.
33. Garbarino E, Slonim R, Sydnor J (2011) Digit ratios (2D:4D) as predictors of risky decision making for both sexes. *J Risk & Unc* 42(1): 1–26.
34. Sapienza P, Zingales L, Maestripieri D (2009) Gender differences in financial risk aversion and career choices are affected by testosterone. *Proc Natl Acad Sci USA* 106 (36): 15268–15273.
35. Breedlove SM (2010) Minireview: Organizational Hypothesis: Instances of the Fingerpost. *Endocrinology* 151(9): 4116-4122.

36. Lutchmaya S, Baron-Cohen S, Raggatt P, Knickmeyer R, Manning JT (2004) 2nd to 4th digit ratios, fetal testosterone and estradiol. *Early Hum Dev* 77(1-2):23-8.
37. Phelps VR (1952) Relative index finger length as a sex-influenced trait in man. *Am J Hum Genet* 4(2): 72–89.
38. Manning JT et al. (2000) The 2nd:4th digit ratio, sexual dimorphism, population differences, and reproductive success: evidence from sexually antagonistic genes. *Evol Hum Behav* 21(3) 163-183.
39. Williams JHG, Greenhalgh KD, Manning JT (2003) Second to fourth finger ratio and developmental psychopathology in preschool children. *Early Hum Dev* 72: 57-65.
40. Bull R, Bensona PJ (2006) Digit ratio (2D:4D) and the spatial representation of magnitude. *Horm Behav* 50(2): 194-199.
41. Poulin M, O'Connell R, Freeman LM (2004) Picture recall skills correlate with 2D:4D ratio in women but not men. *Evol Hum Behav* 25(3): 174-181.
42. Sluming VA, Manning JT (2000) Second to fourth digit ratio in elite musicians: Evidence for musical ability as an honest signal of male fitness. *Evol Hum Behav* 21(1): 1-9.
43. Brookes H, Neave N, Hamilton C, Fink B (2007) Digit ratio (2D:4D) and lateralization for basic numerical quantification. *J Indiv Dif* 28(2): 55-63.
44. Brosnan MJ (2008) Digit ratio as an indicator of numeracy relative to literacy in 7-year-old British school children. *Br J Psychol* 99(Pt 1): 75-85.
45. Austin EJ, Manning JT, McInroy K, Mathews E (2002) A preliminary investigation of the associations between personality, cognitive ability and digit ratio. *Pers Individ Dif* 33(7): 1115–1124.
46. Voracek M, Tran US, Dressler SG (2010) Digit ratio (2D: 4D) and sensation seeking: New data and meta-analysis. *Pers Individ Dif 48*(1): 72-77.
47. Millet K (2009) Low second-to-fourth-digit ratio might predict success among high-frequency financial traders because of a higher need for achievement. *Proc Natl Acad Sci USA* 106(2): 623-628.
48. Lombardo MV et al. (2012) Fetal Programming Effects of Testosterone on the Reward System and Behavioral Approach Tendencies in Humans. *Biol Psychiatry* 72(10):839–847.
49. Coates John (2012) *The Hour Between Dog and Wolf. Risk Taking, Gut Feelings, and the Biology of Boom and Bust,* Penguin Press, New York.
50. Johnson, MW, Bickel, WK (2002) Within-subject comparison of real and hypothetical money rewards in delay discounting. *J Exp Anal Behav* 77(2): 129-146.
51. Lagorio CH, Madden GJ (2005) Delay discounting of real and hypothetical rewards III: Steady-state assessments, forced-choice trials, and all real rewards. *Behav Processes* 69(2): 173-187.
52. Madden GJ, Raiff BR, Lagorio CH, Begotka AM (2004) Delay discounting of potentially real and hypothetical rewards II: between- and within-subject comparisons. *Exp Clin Psychopharmacol* 12(4): 251-261.
53. Coller M, Williams MB (1999) Eliciting individual discount rates. *Exp Econ* 2(2): 107-127.
54. Harrison GW, Lau MI, Williams MB (2002) Estimating individual discount rates in Denmark: A field experiment. *Am Econ Re* 92(5): 1606-1617.
55. Espín AM, Brañas-Garza P, Herrmann B, Gamella JF (2012) Patient and impatient punishers of free-riders. *Proc Biol Sci* 279(1749): 4923-4928.
56. Exadaktylos F, Espín AM, Brañas-Garza P (2013) Experimental subjects are not different. *Sci Rep* 3, 1213, doi: 10.1038/srep01213.
57. Kemper CJ, Schwerdtfeger A (2009) Comparing indirect methods of digit ratio (2D: 4D) measurement. *Am J Hum Biol 21*(2): 188-191.
58. Allaway HC, Bloski TG, Pierson RA, Lujan ME (2009) Digit ratios (2D: 4D) determined by computer-assisted analysis are more reliable than those using physical measurements, photocopies, and printed scans. *Am J Hum Bio 21*(3): 365-370.

APPENDIX

The questions in the tests were asked in Spanish. We provide an English translation.

CRT questions

Spanish:

1. *Un bate y una pelota cuestan 1,10 euros en total. El bate cuesta 1 euro más que la pelota, ¿cuántos céntimos cuesta la pelota?*

2. *Se necesitan 5 máquinas durante 5 minutos para hacer un objeto, ¿cuántos minutos tardarían 100 máquinas en hacer 100 objetos?*

3. *En un lago, hay un conjunto de nenúfares. Cada día, el conjunto se duplica. Si se tardan 48 días en que el conjunto de nenúfares cubra el lago entero, ¿cuántos días tarda el conjunto de nenúfares en cubrir la mitad del lago?*

English:

1. *A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost? _____ cents*

2. *If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? _____ minutes*

3. *In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? _____ days*

Math questions

Spanish:

1. *Si la probabilidad de contraer una enfermedad es de un 10 por ciento, ¿cuántas personas de 1.000 contraerían la enfermedad?*

2. *Si 5 personas tienen el número premiado de la lotería y el premio a repartir es de dos millones de euros, ¿cuánto recibiría cada una?*

3. *Supongamos que tienes 100€ en una cuenta de ahorro, y la tasa de interés que ganas por estos ahorros es de 2% por año. Si mantienes el dinero por 5 años en la cuenta, ¿cuánto tendrá al término de estos 5 años?:*

   a. Más de 102€

   b. Exactamente 102€

   c. Menos de 102€

   d. NS/NR

4. *Digamos que tienes 100€ ahorrados en una cuenta de ahorro. La cuenta acumula un 10% de interés por año. ¿Cuánto tendrás en la cuenta al cabo de dos años?*

English:

1. *If the probability of being infected by an illness is 10%, how many persons of a group of 1000 would be infected by that kind of illness?*

2. *If there are 5 persons that own the winning lottery ticket and the prize to be shared is two million euros, how much money would each person receive?*

3. *Suppose that you have 100€ in a savings account and the rate of interest that you earn from the savings is 2% per year. If you keep the money in the account for 5 years, how much money would you have at the end of these 5 years?:*

   a. More than 102€

   b. 102€ exactly

   c. Less than 102€

   d. S/he cannot/do not want to answer

4. *Suppose that you have 100€ in a savings account. The account accumulates a 10% rate of interest per year. How much money would you have in your account after two years?*

Table A1. Pairwise correlations between variables (by sex)

males	*CRT*	*CRT-1*	*CRT-2*	*CRT-3*	*2D:4D right*	*2D:4D left*	*impatience*
CRT-item 1	0.7101***						
CRT-item 2	0.7346***	0.2575***					
CRT-item 3	0.7712***	0.2903***	0.4090***				
2D:4D right	-0.0630	0.0215	-0.0700	-0.0936			
2D:4D left	-0.0502	-0.0003	-0.0533	-0.0593	0.6580***		
impatience	-0.0201	0.0101	-0.0178	-0.0374	-0.0161	0.0249	
math	0.1258**	0.0405	0.0665	0.1702***	0.0793	0.0530	-0.0743
females							
CRT-item 1	0.7802***						
CRT-item 2	0.6759***	0.2752***					
CRT-item 3	0.7716***	0.3438***	0.3893***				
2D:4D right	-0.1834***	-0.1789***	-0.1602***	-0.0713			
2D:4D left	-0.1322**	-0.0825	-0.1683***	-0.0641	0.7088***		
impatience	-0.1630***	-0.1547***	-0.1035**	-0.0990*	0.0768	0.0253	
math	0.1772***	0.2190***	0.1179**	0.0441	0.0431	0.1114**	-0.0283

Note: Pearson correlations. *,**,*** indicate significance at the 0.1, 0.05 and 0.01 levels, respectively.