Economics Letters 162 (2018) 41–44

Contents lists available at ScienceDirect

Economics Letters

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

Thinking fast, thinking badly

Natalia Jimenez [a,b], Ismael Rodriguez-Lara [b,*], Jean-Robert Tyran [c], Erik Wengström [d]

[a] Departamento de Economia, Metodos Cuantitativos e Historia Economica, Universidad Pablo Olavide, Ctra. de Utrera, 1, Sevilla, E-41013, Spain
[b] Middlesex University London, Department of Economics, Business School , Hendon Campus, The Burroughs, NW4 4BT London, United Kingdom
[c] Department of Economics, University of Vienna, Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria
[d] Department of Economics, University of Lund, S-220 07 Lund, Sweden

HIGHLIGHTS

- We test for the construct validity of the cognitive reflection test.
- Response times indicate that incorrect answers are quicker than correct answers.
- Impulsive subjects complete the test quicker than reflective subjects.
- Our data suggest that intuitive and incorrect answers should be treated differently.

ARTICLE INFO

Article history:
Received 21 June 2017
Received in revised form 12 October 2017
Accepted 20 October 2017
Available online 5 November 2017

JEL classification:
C91

Keywords:
Cognitive abilities
Cognitive reflection
Response times
Intuitive behavior
Reflective behavior

ABSTRACT

We test for the construct validity of the cognitive reflection test (CRT) by eliciting response times. We find that incorrect answers to the CRT are quicker than correct answers. At the individual level, we classify subjects into impulsive and reflective, depending on whether they choose the incorrect intuitive answer or the correct answer the majority of the time. We show that impulsive subjects complete the test quicker than reflective subjects.

1. Introduction

Dual-system models of human thinking differentiate two cognitive processes: a type 1-system that is *fast*, automatic and non-conscious, and a type 2-system that is *slow*, controlled and conscious (Kahneman, 2011; Stanovich and West, 2000). Economists have recently become interested in the relation between these two cognitive processes and decision-making. The cognitive reflection test (CRT) introduced by Frederick (2005) has emerged as a popular tool to identify which way of thinking subjects use. The test consists of three questions that have "*an intuitive answer [that] does spring quickly to mind (…) but this "impulsive" answer is wrong. Anyone who reflects upon it for even a moment would recognize [the correct answer]*" (Frederick, 2005, pp. 26–27).

While scores in the CRT have been related to risk preferences or behavioral biases (Frederick, 2005; Oechssler et al., 2009; Bergman et al., 2010; Hoppe and Kusterer, 2011; Cheung et al., 2014; Brañas-Garza et al., 2012; Andersson et al., 2016), we are not aware of any paper that directly tests the implicit assumption that the CRT measures the tendency to override an intuitive and spontaneous response that is incorrect and to engage in further reflection that leads to giving the correct response. More precisely, we lack evidence about the construct validity of the CRT showing that quick responses to the CRT are likely to be incorrect, while correct answers take longer. Our paper is an attempt to fill out this gap.

2. Data

Hard-copy invitation letters were sent out to a random sample of the Danish population aged between 18 and 80. A total of 2,347 subjects logged on to our webpage and participated (average age = 46.7, SD = 14.3; 1,209 males and 1,138 females). The experiment

* Corresponding author.
E-mail addresses: njimjim@upo.es (N. Jimenez), I.Rodriguez-Lara@mdx.ac.uk (I. Rodriguez-Lara), jean-robert.tyran@univie.ac.at (J.-R. Tyran), erik.wengstrom@nek.lu.se (E. Wengström).

https://doi.org/10.1016/j.econlet.2017.10.018

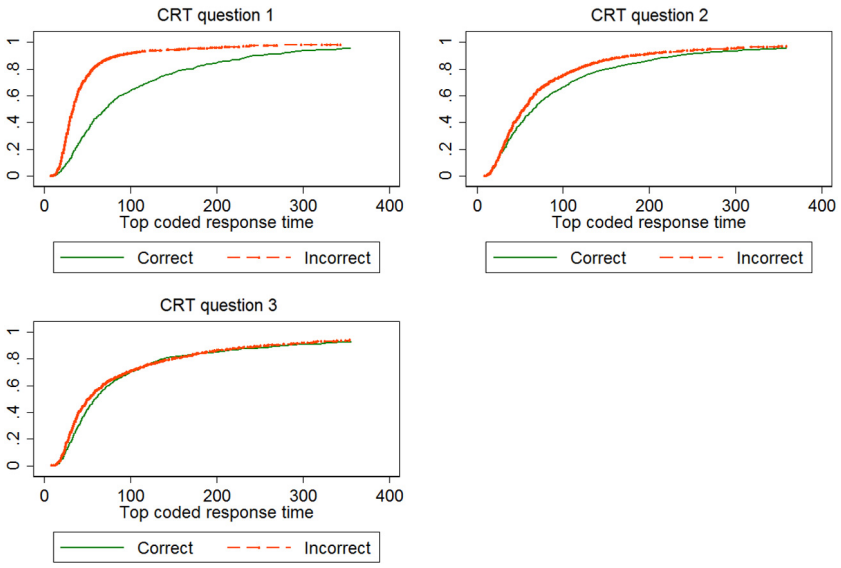

Fig. 1. Cumulative response times to each question.

consisted of two incentivized parts, a public good game (see Thöni et al., 2012; Fosgaard et al., 2014) and a risk elicitation task (see Andersson et al., 2016). The incentivized part was followed by a questionnaire, which included the CRT (Frederick, 2005), as well as basic socio-economic questions, the Big Five personality test and a 20-item cognitive ability test similar to a Raven's progressive matrices test (henceforth referred to as the cognitive ability test).[1]

3. Results

3.1. Correct answers and response times

Fig. 1 depicts the cumulative response times for subjects that gave correct and incorrect answers to each question (see Section A3 in the Online Appendix for more detailed descriptive statistics).[2] We find that subjects who provided the correct answer devoted more time to each question ($p < 0.001$).[3]

Fig. 1 also reveals that the difference in speed between correct and incorrect answers differs across questions. The difference is particularly striking in question 1, and much less pronounced in question 3. It could be that the first question has a more salient intuitive answer, or perhaps subjects figure out after the first question that they need to think longer since these are tricky if not trick questions.[4] Both explanations are consistent with our data since mean response times are increasing with questions (see Section A3 of the Online Appendix). As a robustness check, we collected additional data using an alternative measure of cognitive reflection (Toplak et al., 2014) with randomized and non-randomized questions to test for possible order effects (see Section 4 below). Overall, we do not find evidence of order effects, suggesting that the different patterns observed across questions is likely not due to the order of presentation, but rather due to characteristics of the questions.[5]

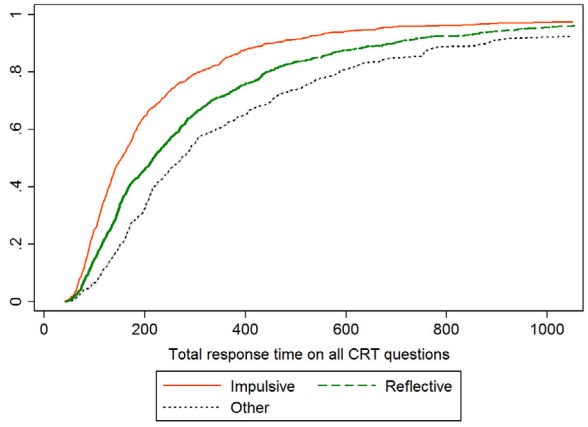

Fig. 2. Cumulative response times of impulsive, reflective and other subjects.

3.2. Intuitive but incorrect answers and response time

Our previous findings support the hypothesis that fast responses are associated with incorrect answers, and vice versa for slow responses. While "impulsive" subjects are frequently defined as those who perform poorly in the CRT, subjects who provide the intuitive (wrong) answer might be treated differently than those who simply provided any incorrect answer (Noussair et al., 2014; Cueva et al., 2015; Ponti and Rodriguez-Lara, 2015). We follow Cueva et al. (2015) and use the iCRT index which adds up the number of intuitive answers, iCRT∈{0,1,2,3}. We then define *Impulsive subjects* as those who scored two or more in the iCRT (39% of the sample) and *Reflective subjects* as those who provided two or more correct answers in the CRT (49% of the sample). The remaining 12% are classified as *Other*.

Fig. 2 displays the cumulative response time distributions for the three types of subjects. We find that Impulsive subjects are

[1] More information about the details of the questionnaire, the recruitment procedures and the sample composition is presented in Sections A1 and A2 of the Online Appendix.

[2] Response times of more than 360 s have been excluded since data contains outliers due to people taking a break or being interrupted. The choice of cut off is not important for any of our results.

[3] Unless otherwise noted, we use the Mann–Whitney and Kolmogorov–Smirnov tests.

[4] These arguments also relate to the "sequence effect" in Brañas-Garza et al. (2015). They report that subjects score better when questions are presented in the standard order, and the smallest (largest) proportion of correct answers is usually observed in question 1 (question 3).

[5] For further details, see Section A6 on the Online Appendix.

Table 1
CRT score, OLS regression.

	(1)	(2)	(3)	(4)
Response time	0.00047***	0.00054***	0.00055***	0.00067***
	[9.10e−05]	[9.07e−05]	[8.91e−05]	[8.52e−05]
Female		−0.422***	−0.406***	−0.366***
		[0.044]	[0.045]	[0.046]
Age 30–39		−0.018	−0.139*	−0.011
		[0.081]	[0.081]	[0.078]
Age 40–49		−0.005	−0.107	0.0422
		[0.072]	[0.073]	[0.071]
Age 50–59		−0.008	−0.086	0.164**
		[0.074]	[0.075]	[0.075]
Age 60–80		−0.241***	−0.333***	0.040
		[0.077]	[0.078]	[0.081]
Basic education			−0.095	−0.048
			[0.081]	[0.077]
Short secondary education			0.164***	0.115**
			[0.056]	[0.054]
Short tertiary education			0.580***	0.461***
			[0.069]	[0.067]
Cognitive ability				0.112***
				[0.007]
Big5 personality scores	No	No	No	Yes
Constant	1.328***[0.034]	1.567*** [0.065]	1.473***[0.065]	0.947***[0.065]
Observations	2,347	2,347	2,347	2,333
R-squared	0.011	0.053	0.088	0.185

Note: Standard errors in brackets. *** p <0.01, ** p <0.05, * p <0.1.

faster (in total response times) than Reflective ones (p <0.001), while Other subjects are slower than both the Impulsive and Reflective ones (p <0.001).

3.3. Regression analysis

To learn more about the relationship between CRT scores and response times, we present a series of regressions in which we control for other factors that are likely to be correlated with both CRT scores and response times.

Table 1 displays the result from a series of OLS regressions using CRT score as the dependent variable. In column 1, we have included response time as the single explanatory variable. The response time has been top-coded at 1080 s, i.e. the response times of subjects that take more than 1080 s are recoded as 1080. Our results are robust to the choice of different cutoffs and also to replacing the OLS with an ordered logit or probit (see Table A6 on the Online Appendix). In columns 2–5, we include additional controls for gender, age, education, cognitive ability and big five personality traits.

The main message of Table 1 is that longer response times are significantly associated with higher CRT scores. A one standard deviation increase in response time corresponds up to a 0.15 standard deviation change in the CRT score. The effect is relatively consistent across specifications and it becomes stronger as we include more control variables. Moreover, CRT scores are related to gender, education and cognitive ability. Notably, there is a negative effect of age effect in columns 2 and 3, but this effect vanishes once we include the cognitive ability test score (column 4). Hence, the decline in CRT with age effect appears to be driven by a decline in cognitive ability.

Table 1 only considers the total score on the CRT and does not distinguish between intuitive and other incorrect answers. Table 2 looks at this issue by means of OLS regressions, where the association to response times is different for Impulsive and Other subjects (the Reflective subjects constitute the left-out category.)[6]

Our estimates show that the relationship reported in Table 1 is driven by the Impulsive subjects, which are on average much faster than the Reflective ones. Impulsive subjects have response times that are up to 0.43 of a standard deviation shorter than the Reflective subjects. To the contrary, subjects who perform poorly on the CRT but are not classified as Impulsive (Other) take more time than the Reflective subjects did. Thus, to measure impulsive thinking, it is important to distinguish between different types of wrong answers and not only count the overall number of incorrect answers. As we include more covariates, the coefficient for the Impulsive subjects increases, whereas the coefficient for the Other subjects decreases. Response times are correlated with gender, age and cognitive ability.

4. Robustness check

We use post-experimental questionnaires in three other experiments to investigate the robustness of our findings. First, we elicited the response time of 311 students (M Age = 21.2 years, SD = 3; 132 males and 179 females) who participated in a laboratory experiment at the Universidad Pablo de Olavide in Seville (Spain). We utilize the alternative measure of cognitive reflection in Toplak et al. (2014) in experiments run at the Universidad de Valencia and the Universidad de Alicante, with a total of 312 participants (M Age = 22.7 years, SD = 5.8; 119 males and 193 females). In addition, Amazon's Mechanical Turk (MTurk) was used to collect data from 195 participants (M Age = 37.8 years, SD = 13.2; 81 males and 114 females) using the Toplak et al.'s version of the test, but with the questions presented in a randomized order.

Overall, we find similar results when we look at these data (a detailed analysis can be found in the Online Appendix, Section A5–A6). We confirm that Impulsive subjects are always faster (in terms of total response time) than Reflective ones (p <0.045). We do not find that the order of presentation affects the difference in response times between correct and incorrect answers in the MTurk experiment.

5. Conclusion

To the best of our knowledge, this paper is the first providing evidence that fast answers to the CRT tend to be incorrect, while

[6] The dependent variable is total response time (in all three questions). Again the results are not sensitive to the choice of cutoff for the top-coding and hold if we instead use a Tobit model or median regressions without top-coding (see Table A7–A8 in the Online Appendix).

Table 2
Total response time, OLS regression.

	(1)	(2)	(3)	(4)
Impulsive	−80.89***	−85.82***	−87.78***	−107.1***
	[10.74]	[10.71]	[10.84]	[11.23]
Other	75.00***	63.12***	60.83***	43.54***
	[16.28]	[16.09]	[16.21]	[16.33]
Female		21.17**	21.24**	18.16*
		[10.04]	[10.27]	[10.98]
Age 30–39		8.955	12.49	6.900
		[17.99]	[18.43]	[18.48]
Age 40–49		22.85	26.20	14.75
		[16.16]	[16.57]	[16.91]
Age 50–59		46.47***	50.48***	26.92
		[16.48]	[16.94]	[17.90]
Age 60–80		128.1***	132.9***	94.68***
		[17.10]	[17.52]	[19.08]
Basic education			−18.16	−19.48
			[18.31]	[18.24]
Short secondary education			−8.822	−7.907
			[12.78]	[12.75]
Short tertiary education			−24.93	−25.53
			[15.93]	[16.11]
Cognitive ability				−11.46***
				[1.73]
Big5 personality scores	No	No	No	Yes
Constant	306.6***	255.8***	263.8***	309.4***
	[7.167]	[14.39]	[15.47]	[62.95]
Observations	2,347	2,347	2,347	2,333
R-squared	0.044	0.077	0.078	0.097

Notes: Standard errors in brackets. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

subjects who take longer tend to provide the correct answer. Our findings lend support to the assumption underlying the test that responses to the CRT can be used to measure the tendency to override intuitive responses. However, our findings also show that merely summing the number of incorrect answers will provide a poor measure of intuitive thinking. Instead, it is important to distinguish between intuitive answers and other types of incorrect answers.

Acknowledgments

We gratefully acknowledge the financial support from the Ministerio de Economia y Competitividad under the projects ECO2016-76789-P (Natalia Jimenez) and ECO2014-58297-R (Ismael Rodriguez-Lara) and the Junta de Andalucia under the project SEJ1436 (Natalia Jimenez), as well as the generous funding provided by The Carlsberg Foundation. Erik Wengström is also thankful for support from Riksbankens Jubileumsfond.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.econlet.2017.10.018.

References

Andersson, O., Holm, H.J., Tyran, J.-R., Wengström, E., 2016. Risk aversion relates to cognitive ability: Preferences or noise? J. Eur. Econ. Assoc. 14 (5), 1129–1154.

Bergman, O., Ellingsen, T., Johannesson, M., Svensson, C., 2010. Anchoring and cognitive ability. Econ. Lett. 107 (1), 66–68.

Brañas-Garza, P., Garcia-Muñoz, T., González, R.H., 2012. Cognitive effort in the beauty contest game. J. Econ. Behav. Organ. 83 (2), 254–260.

Brañas-Garza, P., Kujal, P., Lenkei, B., 2015. Cognitive Reflection Test: Whom, how, when. MPRA Paper No. 68049.

Cheung, S.L., Hedegaard, M., Palan, S., 2014. To see is to believe. Common expectations in experimental asset markets. Eur. Econ. Rev. 66, 84–96.

Cueva, C., Iturbe-Ormaetxe, I., Mata-Perez, E., Ponti, G., Yu, H., Zhukova, V., 2015. Cognitive (ir)reflection: New experimental evidence. J. Behav. Exp. Econ. 64 (5), 81–93.

Fosgaard, T.R., Hansen, L.G., Wengström, E., 2014. Understanding the nature of cooperation variability. J. Public Econ. 120, 134–143.

Frederick, S., 2005. Cognitive reflection and decision making. J. Econ. Perspect. 19 (4), 25–42.

Hoppe, E.I., Kusterer, D.J., 2011. Behavioral biases and cognitive reflection. Econ. Lett. 110 (2), 97–100.

Kahneman, D., 2011. Thinking, fast and slow. Farrar, Straus and Giroux.

Noussair, C.N., Tucker, S.J., Xu, Y., 2014. A futures market reduces bubbles but allows greater profit for more sophisticated traders. University of Waikato, Working Paper in Economics 12/14.

Oechssler, J., Roider, A., Schmitz, P.W., 2009. Cognitive abilities and behavioral biases. J. Econ. Behav. Organ. 72 (1), 147–152.

Ponti, G., Rodriguez-Lara, I., 2015. Social preferences and cognitive reflection: Evidence from a dictator game experiment. Front. Behav. Neurosci. 9 (146).

Stanovich, K.E., West, R.F., 2000. Individual difference in reasoning: Implications for the rationality debate? Behavioural and Brain Sciences 23, 645–726.

Thöni, C., Tyran, J.-R., Wengström, E., 2012. Microfoundations of social capital. J. Public Econ. 96 (7), 635–643.

Toplak, M.E., West, R.F., Stanovich, K.E., 2014. Assessing miserly information processing: An expansion of the cognitive reflection test. Thinking & Reasoning 20 (2), 147–168.