Sex differences on cognitive reflection

Blanca Rosa Olalde*, and Santiago Palacios

University of the Basque Country. UPV/EHU. Donostia-San Sebastian (Spain)

Título: Diferencias sexuales en la reflexión cognitiva.
Resumen: La prueba de reflexión cognitiva o CRT (Cognitive Reflection Test) es una medida popular diseñada para distinguir el procesamiento del pensamiento intuitivo del analítico. Se han encontrado diferencias de sexo en el CRT en múltiples estudios. En esta investigación se estudia la influencia del componente matemático en las diferencias de sexo utilizando dos instrumentos diferentes para medir la reflexión cognitiva: el CRT original y una versión ampliada del original, ambos con un alto componente aritmético y, por otro lado, un conjunto de cinco tareas heurísticas clásicas, denominadas CHT, con un contenido numérico más bajo. Los resultados muestran que estas diferencias de sexo desaparecen cuando se reduce el carácter matemático de las tareas. Finalmente, en este trabajo se aborda una forma más elaborada de puntuar las respuestas intuitivas en relación con el significado de estas diferencias.
Palabras clave: Pensamiento analítico. Reflexión cognitiva. Tarea heurística. Pensamiento intuitivo. Diferencias de sexo.

Abstract: The Cognitive Reflection Test or CRT is a popular measure designed to distinguish the intuitive thinking processing from the analytical one. Sex differences in the CRT have been found in multiple studies. In this research, the influence of the mathematical component on sex differences is studied by using two different instruments to measure cognitive reflection: the original CRT and one extended version of the original one, both with a high arithmetic component and, on the other hand, a set of five classic heuristic tasks, called CHT, with a lower numeracy content. The results show that these sex differences disappear when the mathematical nature of the tasks is reduced. Finally, in this work a more elaborate way of scoring the intuitive responses is addressed in relation to the meaning of these differences.
Keywords: Analytical thinking. Cognitive reflection. Heuristic task. Intuitive thinking. Sex differences.

Introduction

Cognitive reflection is defined as the tendency for individuals to override an intuitive and spontaneous response that turns out to be incorrect; and engage in a more reflective, deliberative and analytical reflection to find the correct answer (Toplak et al., 2014a).

In fact, dual process theories distinguish between the type-1 processes—fast, intuitive, and autonomous—and the type-2 processes—effortful, deliberative and governed by rules. Regarding the framework of the dual process theories, different instruments for measuring the cognitive reflection have emerged over time.

The Cognition Reflection Test created by Frederick (2005) consists of 3 tasks that cue an intuitive but incorrect automatic response. The correct answer requires more deliberation and reflection than it initially seems. The CRT has been widely used and, although it has wide acceptance as a measure of cognitive reflection (Brañas-Garza et al., 2015; Cueva et al., 2016), there are different ways of approaching the cognitive reflection measure (Erceg and Bubić, 2017).

As Toplak et al. (2011) point out, the CRT is well constructed to be used as a predictor of performance in some heuristic tasks. Indeed, a large number of studies (Bialek and Pennycook, 2017; Campitelli and Labollita, 2010; Johnson et al., 2016; Noori, 2016; Shi and An, 2012; Stanovich and West, 2008; Szaszi et al., 2017; Toplak et al., 2011) indicate

that low scores in the CRT are related to several cognitive biases such as overconfidence, the base rate error, the conjunction fallacy, anchoring bias or delay discounting. Recent research has also shown that lower scores are associated with paranormal belief and endorsement of conspiracy theories or fake news (Pennycook and Rand, 2019; Meza et al., 2022).

Significant sex differences have been found (Campitelli and Gerrans, 2014; Frederick, 2005; Toplak, et al., 2011, 2014b) with participants from different age groups, educational level and countries, using both the original CRT and other modified versions (Pennycook et al., 2016; Toplak et al., 2014a). Systematically, men score higher than women in the CRT.

Frederick (2005) suggests that the cause of these sex differences in the CRT is not the lack of attention or effort but the mathematical content of the items; however, he does not confirm the hypothesis. Nevertheless, further investigations take a closer look at this issue. Indeed, although the CRT was originally designed to measure cognitive reflection exclusively, it is evident that it also contains a significant mathematical component (Böckenholt, 2012; Del Missier et al., 2012; Sinayev and Peters, 2015; Welsh et al., 2013). As Toplak (2021) states, cognitive reflection measures require some knowledge, as shown by moderate correlations between cognitive reflection measures and probabilistic numeracy and mathematics (Attali and Bar-Hillel, 2020; Campitelli and Gerrans, 2014; Erceg, et al., 2020; Sinayev and Peters, 2015; Szaszi et al., 2017).

Pennycook and Ross (2016) reported that cognitive reflection, but not numeracy measures, predicted response patterns on moral based dilemmas, suggesting that using a nonnumerical outcome may be more likely to capture variance that is attributable to miserly processing.

---

* Correspondence address [Dirección para correspondencia]:
Blanca Rosa Olalde. Department of Developmental and Educational Psychology, Faculty of Education, Philosophy and Anthropology. University of the Basque Country. UPV/EHU. Donostia-San Sebastian (Spain).
E-mail: blancarosa.olalde@ehu.eus

Due to the numerical content of the items in the CRT, several studies have investigated the relationship between sex differences and mathematical ability (Morsanyi et al., 2018; Primi et al., 2014; Primi et al., 2017; Zhang et al., 2016).

It is necessary to go further and determine whether it is the arithmetic factor involved in the items what contributes to the observed gender differences or it is just the presentation of the tasks with a numerical or mathematical format. In fact, the numerical capacity may be affected both by the perceived self-efficacy in mathematical tests, where women usually obtain lower scores, and by the mathematical anxiety in which women, in general, tend to obtain higher scores (Primi et al., 2016, 2017).

Thus, both the lack of self-confidence in numerical abilities and anxiety towards mathematics can negatively affect the performance in the CRT (Beilock, 2008; Morsanyi et al., 2014; Primi et al., 2016, 2017), but it is expected that men and women do not differ in the level of reflective cognition when controlling these factors.

Although the original CRT is a frequently used measure of cognitive vs. intuitive reflection, some authors claim it is excessively short and emphasize the idea that the CRT is in need of extension (Aczel et al., 2015; Primi et al., 2016, 2017; Thomson and Oppenheimer, 2016; Toplak et al., 2014a). This test may be affected in terms of the floor effect especially in less educated samples. As a result, a large number of researchers consider the use of new instruments that are less dependent on arithmetic or numerical skills to be appropriate.

On the other hand, the recent proposals to score the CRT in different ways are particularly interesting (Erceg and Bubić, 2017; Fuster et al., 2016). Traditionally, all the incorrect answers have been considered as intuitive responses. That fact might prevent from observing the operation of the intuitive thinking correctly. Therefore, it is necessary to implement alternative ways of elaborating the scores in the CRT so that intuitive responses can be clearly differentiated from the incorrect (non-intuitive) ones. That may help to check if sex differences persist as seen so far.

Objectives

The main aim of this study is to check if the sex differences found in the cognitive reflection test (CRT) are related to the mathematical component or the items.

More specific aims of the study are: (1) to confirm sex differences in the sample, (2) to determine if the CRT's mathematical component is linked to the origin of sex differences by using alternative methods to score the answers.

Hypothesis

It is expected to find differences linked to sex in the scores of CRT I (3-items, Frederick, 2005) and CRT II (4-items, Toplak et al., 2014a). The abundant previous literature indicates the persistence of these differences in CRT I. Re-

garding CRT II, these differences are also expected to appear, as the items look similar to the ones of the CRT I in relation to the mathematical presentation.

We expect these differences to diminish or disappear when the tasks lack mathematical presentation or when fewer calculations are required. That is the case with the questionnaire called CHT, consisting of 5 items, adapted from some tasks by Tversky and Kahneman (1974).

Moreover, taking into account just the correct answers linked to reflective reasoning may not allow to see the whole picture. Therefore, it is expected that the original results change significantly when applying other forms of analysing the responses and different criteria for grouping the subjects.

Finally, it is expected that not all items of CRT I, CRT II and CHT show significant sex differences. That fact can help to discern which type of items should be included in a test to measure the cognitive reflection free of the gender gap.

Method

Participants

The sample consisted of 993 students living in the Basque Country (563 females and 430 males), aged 12 - 59 (female: $M = 19$, $SD = 5.913$; male: $M = 19$, $SD = 5.141$). The participants came from nine educational centers and they were performing different studies: students from Secondary Year 9-11 (19.2%), Secondary Year 12-13 (10.9%), Vocational Education and Training (5.1%), Higher Education (4.7%), and University Degree students (51.4%) and Postgraduate students (5.6%). The sample was selected in a non-probabilistic manner.

Instruments

The subjects filled out the CRT I, CRT II and CHT tests in half-hour sessions. Each item of the tests elicited an intuitive answer that was wrong, individuals needed to override that answer and think about the correct answer deliberately. For example, participants were presented with the following item from the CRT I: 'A bat and a ball together cost $1.10. The bat costs $1 more than the ball. How much does the ball cost?' The answer that intuitively and immediately comes to mind is '10 cents', which is incorrect. The correct answer is 5 cents.

The CRT I (Frederick, 2005) contains 3 items: (1) The bat and ball, (2) The machines and time, and (3) The water lily and the lake. The CRT II (Toplak et al., 2014a) includes 4 additional items and it has a new numerical-mathematical format: (4) Barrels and time, (5) Students, (6) Sale results, (7) Results of investment in the stock market.

To make the questionnaire called Classic Heuristic Tasks (CHT), 5 of the classic tasks proposed by Tversky and Kahneman (1974) were selected. All these tasks are related to the representativeness heuristic. They cover (1) the base rate neglect, (2) the sample size, (3) the previous probability, (4)

the conjunction fallacy and (5) the regression to the mean. From these 5 items at least 3 tasks (1, 4 and 5) lack the mathematical format completely.

Analysis of data

In the first step, the traditional methodology that scores the number of correct answers (possible scores 0, 1, 2, 3) was followed in order to study how the way of scoring the answers affects the gender differences. Regarding the incorrect answers, the intuitive answers were not differentiated form other type of incorrect responses. Thus, 2 groups were created and the results were scored as either correct or incorrect. Those subjects with 2 or more correct answers were included in the reflective or analytical group; versus those who gave 2 or more incorrect answers were considered intuitive or impulsive people.

Subsequently, the answers were scored in a different way, by taking into account the incorrect intuitive responses and the incorrect non-intuitive ones separately. In this new way of scoring the answers, three categories were used (logical correct responses, intuitive incorrect responses and non-intuitive incorrect responses).

Regarding the CHT, for the first analysis, the initially selected 5 items were included; whereas for the second analy-sis, 2 items were discarded because they did not allow to distinguish an intuitive response from an incorrect one. Only the base rate bias, the conjunction fallacy and the previous probability task were included in this analysis.

Results

The presentation of the results obtained in the CRT I, CRT II and CHT was done according to these two different ways of scoring answers.

First analysis

Cognition Reflection Test (CRT I)

When following the criterion of the number of correct answers, the general results indicated the existence of sex differences (χ^2 = .000). When analysing the number of correct answers according to the sex (See Table 1), it was observed that there were more females than males who did not give any correct answer in the CRT I. In addition, males had more correct answers than females. On average, males resolved .55 of the three items in the CRT 1 (SD = 0.81) compared to .30 (SD = 0.62) among the females.

Table 1
Percentage of correct answers.

Number of correct answers	CRT I			CRT II			CHT		
	M	F	Total	M	F	Total	M	F	Total
0	62.7	77.5	71.1	31.6	36.5	34.3	12.4	11.0	11.6
1	21.9	16.1	18.6	39.6	42.5	41.2	24.0	35.4	30.4
2	12.7	5.0	8.3	17.3	15.8	16.4	36.4	29.4	32.4
3	2.8	1.4	2.0	8.5	4.6	6.3	19.8	18.8	19.2
4				3.0	.7	1.7	3.2	3.2	3.2
5							0.0	0.0	0.0

Note: M: Male; F: Female.

As hypothesized, the measure of the reflective thinking obtained through the number of correct answers in CRT I did significantly correlate with sex, applying both parametric (P .000) and nonparametric techniques. Thus, both in the Mann-Whitney U test and the Kruskal-Wallis test, the level of significance was .000. This indicated that both groups had a different performance with respect to this measure of cognitive reflection.

In relation to the 2-group classification (reflexive vs. impulsive), the differences still remained significant (χ^2 = .005).

When each item was addressed (see Table 2), the results indicated that while there were no differences in the first item (The bat and the ball), differences could be observed in the second (Machines and time) and the third (The water lily and the lake) items. Thus, despite the difference in the percentages of the correct responses by males (11.5%) and females (9.5) in item 1, the difference was not significant (Pearson's χ^2 = 300).

According to the previous literature (Frederick, 2005), the most difficult problem to be solved was the first item and it was likely that this was the origin of some difficulties to detect the differences, which were statistically significant in items 2 and 3 (Pearson's χ^2 = .000, in both cases).

Cognition Reflection Test (CRT II)

On average, males solved 1.13 from the four items of the CRT II correctly (SD = 1.04), females solved fewer (M = 0.9, SD = 0.87) (see Table 1) and according to the classification of subjects in 2 groups, reflexive vs impulsive, differences were again significant (χ^2 =0.005).

As suggested in the hypothesis, the results obtained in the CRT I were also confirmed in the CRT II and there were significant sex differences (P .001). Thus, both in the Mann-Whitney U test and the Kruskal-Wallis test, the level of significance obtained was 5.000 respectively in the CRT II. Our

data indicated that both groups had different performance in this case as well.

On analyzing the 4 items of the CRT II individually, the results (see Table 2) indicated that only in the item 1 (Barrel time problem) the differences were really significant and especially items 2 and 3 stood out because they were far from showing those differences.

Table 2
Percentage of subjects with correct answers in each item.

CRT I	M	F	Total	χ^2 Pearson
Item 1	11.5	9.5	10.4	0.300
Item 2	21.2	9.5	14.6	0.000
Item 3	24.0	12.8	17.8	0.000
CRT II				
Item 1	21.9	9	14.5	0.000
Item 2	12.2	10.4	11.2	0.378
Item 3	40	38.7	39.3	0.694
Item 4	38.6	33.4	25.6	0.089
CHT				
Item 1	3.7	3.2	3.4	0.665
Item 2	58.8	46.4	51.8	0.000
Item 3	24.0	21.6	22.6	0.361
Item 4	43.5	38.6	40.7	0.113
Item 5	46.3	57.7	52.8	0.000

Note: M: Male; F: Female.

Classic Heuristic Tasks (CHT)

The proposed hypothesis was confirmed and the CHT scores did not show this bias linked to gender, similar scores were obtained in both groups (P, 82 in the Mann- Witney U test). When analysing the items individually, the results (see Table 2) were relevant since they indicated that the scores of items 1 (Base rate), item 3 (Previous probability) and item 4 (Conjunction fallacy) were not linked to sex. On the other hand, in item 2 (Sample size) and 5 (Regression to the mean) differences appeared, but in an opposite direction. Thus, in item 2, men had the best performance, while in item 5 the findings revealed that women performed better.

So far it has been observed that the CRT I and II are psychometrically as expected (including the presence of bias linked to sex) and that result confirms the validity and reliability of the sample studied to a large extent.

On the other hand, CHT happened to be more sensitive and less difficult than the CRT I and II. However, no subject reached the maximum score, that is 5, due to the difficulty of the item 1 (Base Rate), which was solved only by 3.4% of the participants correctly. In this case, similarly to item 1 in the CRT I, the sex differences might have probably disappeared as a consequence of the difficulty of the item itself.

Second analysis

Intuitive answers and other incorrect responses

To verify whether the sex differences were maintained or disappeared when the answers were analysed according to their genuinely intuitive nature, the responses were classified in three categories: logical correct responses, intuitive incorrect responses and non-intuitive incorrect ones.

Table 3
Mean scores by men and women in the answers of the three tests.

Responses	CRT I			CRT II			CHT (3)		
	H	M	Sig.*	H	M	Sig.	H	M	Sig.
Logical	0.56	0.30	0.0	1.12	0.91	0.008	0.70	0.62	0.112
Intuitive	1.44	1.59	0.0	1.69	1.80	0.200	1.12	1.21	0.053
Incorrect	0.94	1.08	0.022	1.19	1.29	0.034	1.17	1.16	0.663

Note: * Mann-Whitney U test.

With regard to the CRT I, the results showed that, although differences continued to persist, they affected both the logical responses (Sig. 000 in Man-Whitney U) and the intuitive ones (Sig. 3.000), but not the incorrect non-intuitive answers, where the differences were not significant (Sig. 57.000) between males and females.

As for the CRT II, the differences affected the logical responses (Sig. 5.000) and the incorrect non-intuitive answers (Sig. 48.000) but not the intuitive responses (Sig. 154.000).

When analysing each of the CRT I items according to this elaboration of the answer, the results were similar to the first analysis, where items 2 and 3 showed significant sex differences. Regarding to the CRT II, the results of the research indicated that item 1 was again responsible for the differences.

Finally, in the CHT there were no sex differences in any of the three types of responses when looking into the individual analysis of each of the items (see Table 4), according to this criterion, the analysis did not reveal any sex differences in any of the 3 items.

Table 4
Percentage of responses in each item of the CHT.

Responses	Item 1			Item 2			Item 3		
	M	F	Sig.	M	F	Sig.	M	F	Sig.
Logical	3.5	3.2		23.5	21.2		42.8	37.4	
Intuitive	65.4	69.9	0.334	7.1	5.8	0.437	40.5	46.7	0.141
Incorrect	13.7	15.1		69.4	73.0		16.7	16.0	

In conclusion, we found that the CRT I was the most biased questionnaire and that CHT was the less biased one, although the latter was not totally free of it (Items 2 and 5). Moreover, an analysis of the correlations was carried out (see Table 5) to estimate if the CHT measured the same cognitive reflection construct as the CRT I and

II. The results indicated that there was a positive correlation between all the CHT, CRT I and II logical responses. In addition, the CHT logical responses had negative correla-

tions with the heuristic responses of both the CRT I and the CRT II.

On the other hand, the intuitive CHT responses did not show significant correlations with those of CRT I and II, while those of the CRT I and II showed significant positive correlations with each other (0.201 **).

Finally, the incorrect CHT responses showed positive correlations with the incorrect CRT I responses, but not with CRT II responses. In addition, they showed negative correlations with the logical responses of both the CRT I and II.

Although the correlations were not statistically significant in this sample, in other studies (Olalde, 2021) the intuitive responses of the CHT showed positive correlations with the intuitive responses of the CRT I but not with those of the CRT II. Likewise, the incorrect CHT responses showed positive correlations with the intuitive responses of the CRT II.

Table 5
Correlations between CHT and CRT I and II scores.

		CRT I			CRT II		
		Log.R.	Int. R.	Inc. R.	Log. R	Int. R.	Inc. R.
CHT	Pearson Correlation Sig. (2-tailed) Log. R	0.121* *	-0.065*	-0.044	0.104* *	-0.076*	-0.023
	Int. R.	-0.026 - 0.093* *	0.062 0.008	-0.047 0.083**	-0.027 - 0.083* *	-0.005 0.062	0.037 0.015
	Inc. R.						

Note: Log. R.: Logical Responses; Int. R.: Intuitive Responses; Inc. R.: Incorrect Responses

In order to delve into these relations, a regression analysis was carried out, the score of the logical responses in the CRT I was viewed as the dependent variable and the logical scores in the CRT II and in the CHT were viewed as independent variables. When performing the analysis with the complete sample, both the CRT II and the CHT explained (.000 and .007 respectively) the CRT I scores significantly. In particular, the CRT II explained 41.5% of the variance and the CHT the 7.9%.

However, when performing this regression analysis for males and females separately (see Table 6), it was observed that while in the case of females the CRT II explained 33.6% of the variance, the CHT did not reach to be significant, explaining scarcely the 2.7% of the variance. In the case of males, both the CHT and the CRT II helped to explain more than 55% (11.4% and 46% respectively) of the variance in the logical scores of the CRT I.

Table 6
Results of the regression analysis across gender.

	B	S E	Beta	*t*	*p*
Females					
CRT II	0.305	0.038	0.319	7.961	.0
CHT	0.034	0.039	0.035	0.874	.383
Males					
CRT II	0.485	0.045	0.464	10.893	.0
CHT	0.115	0.049	0.101	2.373	.018

Conclusions

The creation of a cognitive reflection measure that may be less influenced by the sex bias is needed. Reducing the mathematical nature of the items and introducing new ways of analysing the answers is recommended.

As hypothesized, the sex differences were evident both in the CRT I and in the CRT II, although in the latter they became diluted by differentiating the incorrect answers from the intuitive one. Our results in the CHT proved that the mathematical nature of the items can affect the scores obtained by means of the CRT. Thus, when the numerical presentation of the items decreased, the sex differences disappeared. On the other hand, there was an acceptable degree of correlation between the scores of the 3 instruments used in the research. That allowed us to think about the possibility of measuring cognitive reflection through tasks without this mathematical aspect.

It was also observed that not all the items contributed to these differences equally. For this reason, it is advisable to expand the measurement of the cognitive reflection, including items that allow to observe the process associated with it; such as intuitive response detection, rejection of that response, the start-up of the type-2 processes and the search and achievement of a better response (Bialek and Pennycook, 2017), as well as the deletion of sex biases.

We have provided further evidence that the sex differences can be significantly reduced when the effects of numeracy are controlled (Primi et al., 2016).

On the other hand, it is especially interesting to be able to observe the behavior of the properly intuitive responses and not only those that point to the use of logical processing. We also propose that further research designs for the analysis of the data should be carried out to detect the subjects who do not even activate a genuinely intuitive response.

The development of different scores can help to clarify the differences, as they allow more meaningful groupings of the subjects considering not only the logical responses, but also the intuitive ones (Erceg and Bubić, 2017). Some protocol analysis to unfold the steps of the reasoning process in solving the CRT is recommendable (Szaszi et al., 2017), because there are several ways people solve or fail the test.

Apart from studying the effects of numeracy, it is clear that future research should focus on other variables related to sex differences. The stereotype related to the mathematical performance in women, the mathematical anxiety, the perceived self- efficacy in mathematical tests or the gender different cognitive responses to competitive pressure in standardized tests (Beilock, 2008; Morsanyi et al.., 2014; Primi et al., 2016, 2017) can influence the gender differences in the measure of cognitive reflection through the CRT. In short, it is necessary to have new instruments to measure cognitive reflection free of the sex differences (Stanovich et al., 2016).

The conclusions of this study must be applicable to the educational sphere. That should guarantee the equal development of skills by all students, regardless of their gender, origin or any other personal characteristic, in order to ensure the improvement of opportunities in society. Čavojová and Hanák (2016) highlight the need for a better education on cognitive reflection and the need to teach people how to reflect on their beliefs and intuitions.

To sum up, this work aims to overcome gender stereotypes related to intuitive and analytical thinking and to avoid identifying women with intuitive thinking and men with analytical thinking. It is important not to think of intuitive and analytical thinkers as two different types of people, since we are all capable of both types of reasoning.

Everybody can think analytically under the right circumstances.

References

Aczel, B., Bago, B., Szollosi, A., Foldes, A., & Lukacs, B. (2015). Measuring individual differences in decision biases: Methodological considerations. *Frontiers in Psychology, 6*, 1770. https://doi.org/10.3389/fpsyg.2015.01770

Attali, Y., & Bar-Hillel, M. (2020). The false allure of fast lures. *Judgment & Decision Making, 15*(1). https://doi.org/10.1017/S1930297500006938

Beilock, S. L. (2008). Math performance in stressful situations. *Current Directions in Psychological Science, 17*(5), 339-343.

Bialek, M., & Pennycook, G. (2017). The cognitive reflection test is robust to multiple exposures. *Behavior Research Methods, 50*(5), 1953–1959. https://doi.org/10.3758/s13428-017-0963-x

Böckenholt, U. (2012). The cognitive-miser response model: Testing for intuitive and deliberate reasoning. *Psychometrika, 77*(2), 388-399. 10.1007/s11336-012-9251

Brañas-Garza, P., Kujal, P., & Lenkei, B. (2015). *Cognitive reflection test: Whom, how, when.* Economic Science Inst.

Campitelli, G., & Gerrans, P. (2014). Does the cognitive reflection test measure cognitive reflection? A mathematical modeling approach. *Memory & Cognition, 42*(3), 434-447. https://doi.org/10.3758/s13421-013-0367-9

Campitelli, G., & Labollita, M. (2010). Correlations of cognitive reflection with judgments and choices. *Judgment and Decision Making, 5*(3), 182-191.

Čavojová, V., & Hanák, R. (2016). Cultures influences on cognitive reflection. In J. V. Cohn, S. Schatz, H. Freeman, and D. J. Y. Combs (Eds.), *Modeling Sociocultural Influences on Decision Making: Understanding Conflict, Enabling Stability* (pp. 85-102). Routledge.

Cueva, C., Iturbe-Ormaetxe, I., Mata-Perez, E., Ponti, G., Sartarelli, M., & Yu, H. (2016). Cognitive (ir)reflection: New experimental evidence. *Journal of Behavioral and Experimental Economics, 64*, 81-93. https://doi.org/0.1016/j.socec.2015.09.002

Del Missier, F., Mäntylä, T., & Bruin, W. B. (2012). Decision-making competence, executive functioning, and general cognitive abilities. *Journal of Behavioral Decision Making, 25*(4), 331-351. https://doi.org/10.1002/bdm.731

Erceg, N., & Bubić, A. (2017). One test, five scoring procedures: Different ways of approaching the cognitive reflection test. *Journal of Cognitive Psychology, 29*(3), 381-392. https://doi.org/10.1080/20445911.2016.1278004

Erceg, N., Galić, Z., & Ružojčić, M. (2020). A reflection on cognitive reflection-testing convergent/divergent validity of two measures of cognitive reflection. *Judgment & Decision Making, 15*(5). https://doi.org/10.1017/S1930297500007907

Frederick, S. (2005). Cognitive reflection and decision making. *The Journal of Economic Perspectives, 19*(4), 25-42. https://doi.org/10.1257/089533005775196732

Fuster, B., Agulló Candela, J., Ferreira Magalhaes, M., Fuster, A., Kostova, M., & Sartarelli, M. (2016). Habilidad Cognitiva y Sistemas de Evaluación. En M. T. Tortosa Ybáñez, S. Grau Company, and J. D. Álvarez Teruel (Coords.), *XIV Jornadas de Redes de Investigación en Docencia Universitaria: investigación, innovación y enseñanza universitaria: enfoques pluridisciplinares* (pp. 2580-2593)

Johnson, E. D., Tubau, E., & De Neys, W. (2016). The doubting system 1: Evidence for automatic substitution sensitivity. *Acta Psychologica, 164*, 56-64. https://doi.org/10.1016/j.actpsy.2015.12.008

Meza, J. M., Aranda, C. E, Olalde, B. R., Palacios, S. (2022). Sesgos cognitivos y la discriminación de noticias falsas en una muestra de adultos mexicanos. *Revista de Psicología de la Universidad Autónoma del Estado de México, 11* (27)*, 2.*

Morsanyi, K., Busdraghi, C., & Primi, C. (2014). Mathematical anxiety is linked to reduced cognitive reflection: A potential road from discomfort in the mathematics classroom to susceptibility to biases. *Behavioral and Brain Functions, 10*(1), 31. https://doi.org/10.1186/1744-9081-10-31

Morsanyi, K., Prado, J., & Richland, L. E. (2018). Editorial: The role of reasoning in mathematical thinking. *Thinking & Reasoning, 24*(2), 1-9. https://doi.org/10.1080/13546783.2018.1435425

Noori, M. (2016). Cognitive reflection as a predictor of susceptibility to behavioral anomalies. *Judgment and Decision Making, 11*(1), 114–120. https://doi.org/10.1017/S1930297500007634.

Olalde, B. R. (2021). *Procesamiento dual y conocimiento social. Interacción del razonamiento lógico-heurístico y las creencias sobre los estereotipos sociales del colectivo LGTBI y la comunidad gitana.* Universidad del País Vasco / Euskal Herriko Unibertsitatea, Argitalpen Zerbitzua.

Pennycook, G., Cheyne, J., Koehler, D., & Fugelsang, J. (2016). Is the cognitive reflection test a measure of both reflection and intuition? *Behavior Research Methods, 48*(1), 341-348. https://doi.org/10.3758/s13428-015-0576-1

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015). Everyday consequences of analytic thinking. *Current Directions in Psychological Science, 24*(6), 425-432. https://doi.org/10.1177/0963721415604610

Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition, 188*, 39-50. https://doi.org/10.1016/j.cognition.2018.06.011

Primi, C., Donati, M. A., Chiesi, F., & Morsanyi, K. (2017). Are there gender differences in cognitive reflection? Invariance and differences related to mathematics. *Thinking & Reasoning, 36*(5), 1-22. https://doi.org/10.1080/13546783.2017.1387606

Primi, C., Morsanyi, K., Chiesi, F., & Donati, M. A. (2014). Item response theory analysis of the cognitive reflection test: Testing the psychometric properties of the original scale and a newly developed 8-item version. Paper presented at the 2799- 2804. https://doi.org/10.1002/bdm.1883

Primi, C., Morsanyi, K., Chiesi, F., Donati, M. A., & Hamilton, J. (2016). The development and testing of a new version of the cognitive reflection test applying item response theory (IRT). *Journal of Behavioral Decision Making, 29*(5), 453- 469. https://doi.org/10.1002/bdm.1883

Shi, X., & An, N. (2012). Cognitive ability and psychological biases: Perspective from chinese stock individual investors. https://www.diva-portal.org/smash/get/diva2:513883/FULLTEXT01.pdf

Sinayev, A., & Peters, E. (2015). Cognitive reflection vs. calculation in decision making. *Frontiers in Psychology, 6,* 532. https://doi.org/10.3389/fpsyg.2015.00532

Stanovich, K. E., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology, 94*(4), 672. https://doi.org/10.1037/0022-3514.94.4.672

Stanovich, K. E., West, R. F., & Toplak, M. E. (2016). Toward a rationality quotient (rq). *The Thinking Mind: A Festschrift for Ken Manktelow,* 202. http://keithstanovich.com/Site/Research_on_Reasoning_files/Stanovich%20in%20Galbraith_proofs.pdf

Szaszi, B., Szollosi, A., Palfi, B., & Aczel, B. (2017). The cognitive reflection test revisited: Exploring the ways individuals solve the test. *Thinking &*

Reasoning, *23*(3), 207-234. https://doi.org/10.1080/13546783.2017.1292954

Thomson, K. S., & Oppenheimer, D. M. (2016). Investigating an alternate form of the cognitive reflection test. *Judgment and Decision Making, 11*(1), 99-113. https://doi.org/10.1017/S1930297500007622

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The cognitive reflection test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition, 39*(7), 1275-1289. https://doi.org/10.3758/s13421-011-0104-1

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014a). Assessing miserly information processing: An expansion of the cognitive reflection test. *Thinking & Reasoning, 20*(2), 147-168. https://doi.org/10.1080/13546783.2013.844729

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014b). Rational thinking and cognitive sophistication: Development, cognitive abilities, and thinking dispositions. *Developmental Psychology, 50*(4), 1037-1048. https://doi.org/10.1037/a0034910

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124-1131.

Zhang, D. C., Highhouse, S., & Rada, T. B. (2016). Explaining sex differences on the Cognitive Reflection Test. *Personality and Individual Differences*, *101*, 425-427. https://doi.org/10.1016/j.paid.2016.06.034