



Criterion Validity of Cognitive Reflection for Predicting Job Performance and Training Proficiency: A Meta-Analysis

Inmaculada Otero*, Jesús F. Salgado and Silvia Moscoso

Department of Political Science and Sociology, Faculty of Labor Relations, University of Santiago de Compostela, Santiago de Compostela, Spain

This article presents a meta-analysis of the validity of cognitive reflection (CR) for predicting job performance and training proficiency. It also examines the incremental validity of CR over cognitive intelligence (CI) for predicting these two occupational criteria. CR proved to be an excellent predictor of job performance and training proficiency, and the magnitude of the true validity was very similar across the two criteria. Results also showed that the type of CR is not a moderator of CR validity. We also found that CR showed incremental variance over CI for the explanation of job performance, although the magnitude of the contribution is small. However, CR shows practically no incremental validity over CI validity in the explanation of training proficiency. Finally, we discuss the implications of these findings for the research and practice of personnel selection.

Keywords: cognitive reflection, job performance, training proficiency, meta-analysis, cognitive intelligence

OPEN ACCESS

Edited by:

Jungkun Park,
Hanyang University, South Korea

Reviewed by:

Isabel Mercader Rubio,
University of Almería, Spain
Ioannis Anyfantis,
European Agency for Occupational
Safety and Health, Spain

*Correspondence:

Inmaculada Otero
inmaculada.otero@usc.es

Specialty section:

This article was submitted to
Organizational Psychology,
a section of the journal
Frontiers in Psychology

Received: 16 February 2021

Accepted: 22 April 2021

Published: 31 May 2021

Citation:

Otero I, Salgado JF and
Moscoso S (2021) Criterion Validity
of Cognitive Reflection for Predicting
Job Performance and Training
Proficiency: A Meta-Analysis.
Front. Psychol. 12:668592.
doi: 10.3389/fpsyg.2021.668592

INTRODUCTION

Cognitive reflection (CR) is one of the most well-known concepts and is based on the view that the human mind operates through two types of cognitive processes known as System 1 (S1) and System 2 (S2). There is high consensus among researchers (Wason and Evans, 1975; Evans and Wason, 1976; Chaiken et al., 1989; Metcalfe and Mischel, 1999; Kahneman and Frederick, 2002, 2005; Stanovich and West, 2002; Strack and Deutsch, 2004; Kahneman, 2011) that S1 is fast, automatic, unconscious and impulsive and operates with little (or no) effort, and that S2 is slow, reflective, and purposeful and requires effort and concentration. Moreover, S2 requires attention, and it is capable of solving complex problems with a high degree of accuracy. On the contrary, S1 tends to use heuristics, biases, and shortcuts to operate quickly and without effort.

Kahneman and Frederick (2002, 2005), Frederick (2005), and Kahneman (2011) defined cognitive reflection as the individual capacity to annul the first impulsive response that the mind offers, and to activate the reflective mechanisms that allow to find a response, make a decision, or carry out a specific behavior. The first impulsive answer is frequently wrong, and cognitive reflection would activate a more reflective and correct answer.

Kahneman and Frederick (2002, 2005), Frederick (2005), and Kahneman (2011) developed the famous cognitive reflection test (CRT) to evaluate individual differences in cognitive

reflection. This test consists of three apparently simple arithmetical problems with two alternative answers. Each problem elicits a quick but wrong response. The second answer requires effort, deliberation, and reflection and is the right one. If individuals respond with the first-and-wrong response, they will score low in cognitive reflection. On the contrary, if individuals suppress this first-and-wrong answer in favor of an alternative one, they would be high in cognitive reflection. This test is currently very popular. For example, a search in Google indicates that there are over 38,000 entries with the label “Cognitive Reflection Test,” and Wikipedia also has an entry devoted to the CRT.

Although the CRT (also called 3-item CRT) is the most famous test, other forms of the CRT were developed over the last decades. For instance, some researchers have developed larger CRTs (see, Salgado et al., 2019; Sirota et al., 2020), and others have added new items to the CRT-3 items (e.g., Finucane and Gullion, 2010; Toplak et al., 2014; Primi et al., 2015). Recently, some researchers have shown an interest in the numerical content of the CRT, and they have developed verbal-CRTs (e.g., Thomson and Oppenheimer, 2016; Sirota et al., 2020). These tests also trigger an immediate answer as in the 3-item CRT, and they might involve numbers in their statements, but mathematical calculations are not required in order to find the correct answer.

The relationship between the CRTs and significant cognitive and social criteria has been examined. For instance, empirical research showed that CR predicts achievement in hypothetical decision-making tasks, including intertemporal choice tasks and risky choice tasks. The results suggested that people who score higher in CR showed a preference for delayed-but-bigger gratification against an immediate-but-smaller one (Frederick, 2005; Oechssler et al., 2009). Research also showed that people with higher CR scores tended to display more riskier behaviors in hypothetical financial gain scenarios, even when the expected value of the risky option was lower than the expected value of the safe option (Frederick, 2005). Studies carried out to explore economic behavior employing game theory have suggested that people who scored higher in CR tend to give more accurate answers, deviating less from the normative answers and, consequently, obtaining better results (Moritz et al., 2013; Baghestanian et al., 2015; Corgnet et al., 2015a; Georganas et al., 2015; Noussair et al., 2016; Kocher et al., 2019). It was also found that CR predicted heuristic behavior in decision-making tasks and judgment tasks. For instance, CR predicted the avoidance of several cognitive biases (e.g., base rate neglect, sunk cost effect, conjunction fallacy, and anchoring effect, among others; see for instance, Campitelli and Labollita, 2010; Toplak et al., 2011, 2014) and the resistance to stereotypes and prejudices (Lubian and Untertrifaller, 2013) while making judgments and decisions.

In the organizational domain, the relationships between CR and job performance and training proficiency have also been studied. However, the empirical studies offer a variety of results, some of them showing no relationship between CR and job performance (e.g., Lado et al., 2021) and others showing a moderate correlation (e.g., Corgnet et al., 2015b, 2019; Otero, 2019; Salgado et al., 2019). Research on the relationship between

CR and training proficiency also offers an optimistic view, as many studies reported that CR was a relevant predictor of this criterion, although there was also observed variability in the estimates of the relationship (Toplak et al., 2014; Gómez-Veiga et al., 2018; Otero, 2019; Muñoz-Murillo et al., 2020).

When the empirical studies show considerable variability in the relationship between two variables and it is difficult to reach a consensus, meta-analytical techniques offer a potential way of solving the dispute. However, no previous research has meta-analytically examined the relationship between CR and these two critical organizational criteria. Consequently, this article aims to shed further light on the relationships between CR and job performance and training proficiency by providing an estimate of the average correlation between these variables and by examining whether the observed variability is real or artifactual. Moreover, this study also aims to determine whether the type of the CR test moderates the relationship between CR and job performance and training proficiency. Finally, the study also aims to explore whether CR adds validity over cognitive intelligence (i.e., the best predictor of job performance and training proficiency) for predicting both organizational criteria. These four issues are the main goals of this study.

CR, Job Performance, and Training Proficiency

According to Campbell et al. (1990) and Campbell and Wiernik (2015), job performance refers to any cognitive, psychomotor, motor, or interpersonal behavior, which is controlled by individual, gradable in terms of ability, and relevant for organization goals. Supervisor ratings, work sample test, assessment center ratings, and production records are examples of job performance measures. To our knowledge, few studies have explored the validity of CRT as a predictor of job performance (Morsanyi et al., 2014; Corgnet et al., 2015b; Thoma et al., 2015; Čavojská and Jurkovič, 2017; Otero, 2019; Salgado et al., 2019; Lado et al., 2021). The study carried out by Corgnet et al. (2015b) used a laboratory task as a proxy of job performance. They found that CRT significantly predicted job performance. The study of Čavojská and Jurkovič (2017) used two CRTs, Frederick's 3-item CRT and a 7-item version. They showed a significant correlation with teacher proficiency. The study of Morsanyi et al. (2014) reported that CR predicted job performance as measured in a laboratory task. Salgado et al. (2019) used an overall rating of an assessment center as the measure of job performance, and Otero (2019) used a work sample test to assess job performance. Both studies found that CR was a predictor of these measures of job performance. Otero (2019) and Salgado et al. (2019) also reported substantial observed validities. More recently, Lado et al. (2021) reported that CR did not significantly correlate with overall job performance ratings. In summary, the literature on CR validity for predicting job performance is relatively scarce, uses two types of CRT (standard version and enlarged versions), and a variety of ways of assessing job performance (e.g., laboratory tasks, assessment center ratings, and work sample tests). The correlations between CR and job performance ranged from 0.05 to 0.47.

Training proficiency refers to the degree of technical skill and competence acquired after a period of education or instruction. Grades, marks, and instructor ratings are examples of training proficiency measures. Previous studies have shown that CR is a relevant predictor of several measures of training proficiency. For example, Muñoz-Murillo et al. (2020) found that the CRT substantially correlated with grade point average (GPA), and Toplak et al. (2014) also reported a significant validity for predicting self-reported GPA. Thomson and Oppenheimer (2016) found that a verbal CRT and a numerical CRT were valid predictors of GPA. Other studies conducted by Morsanyi et al. (2014); Otero (2019), and Salgado et al. (2019) also found substantial correlations. However, Corgnet et al. (2015b) did not find correlation between CRT and GPA.

Other indicators of training proficiency have also been explored. For example, research has shown that CRT was a predictor of mathematical achievement (Gómez-Chacón et al., 2014; Gómez-Veiga et al., 2018), physics qualifications (Lindeman and Svedholm-Häkkinen, 2016), scores in the math and language section of the German A-level exam (Lohse, 2016), statistics exam qualifications (Primi et al., 2017), and educational level (Yilmaz and Sarıbay, 2016, 2017; Skagerlund et al., 2018). However, CRT was not a predictor of school attendance (Primi et al., 2017).

Despite the fact that the number of studies that have examined the relationship between CR and job performance and training proficiency is relatively low, the majority of the studies found that the CR test predicts both organizational criteria. However, there is substantial variability in the magnitude of the correlations. There are both methodological and theoretical reasons that might explain the variability. Artifacts (i.e., sampling error, measurement error, and range restriction) might explain a part of or the totality of the observed variability. Another source of variation in the correlations might be that the studies used different CRTs, with different reliabilities. It was repeatedly found that the CRT-3 showed low internal consistency and that the enlarged version of the CRT and the new CRTs have higher reliability than the original one (see Otero, 2019; Salgado et al., 2019; Otero et al., 2021 for more details). Reliability (measurement error) produces two artifactual effects. First, it attenuates the magnitude of the correlations. Second, it introduces error variance into the distributions of correlations. The third source of variation might be that the studies used a variety of estimates of job performance and training proficiency, which have different psychometric properties (e.g., construct validity and reliability). Based on the research discussed above, we posit the following research questions:

Research Question 1: What is the correlation between CR and job performance?

Research Question 2: What is the correlation between CR and training proficiency?

Research Question 3: Is the observed variability in the correlation between CR and job performance and training proficiency real or artifactual?

Research Question 4: Does the type of CRT moderate the relationship between CR and job performance and training proficiency?

Incremental Validity of CR

Some studies have examined the incremental validity of CR over other variables for predicting training proficiency and job performance, although the number of studies is small. Cognitive intelligence (CI) occupies a special place in these sorts of studies because many meta-analyses have indicated that CI is the best predictor of job performance and training proficiency and that CI generalizes the validity across organizations, jobs, and samples (e.g., Hunter and Hunter, 1984; Kuncel et al., 2001, 2005; Salgado et al., 2003; Bertua et al., 2005; Richardson et al., 2012; Salgado and Moscoso, 2019). For instance, a recent meta-analysis conducted by Salgado and Moscoso (2019) found that CI showed an average true validity of 0.44 and 0.62 for predicting job performance and training proficiency, respectively.

Two studies (Otero, 2019; Salgado et al., 2019) tested if CR showed incremental validity of CR over CI for predicting job performance and training proficiency. Salgado et al. (2019) found that the joint validity of CR and CI for predicting job performance was higher than the validity of CI alone, but the incremental validity was small ($\Delta R^2 = 0.03$). Likewise, Otero (2019) found that CR added validity over CI, but again the incremental validity was small ($\Delta R^2 = 0.01$).

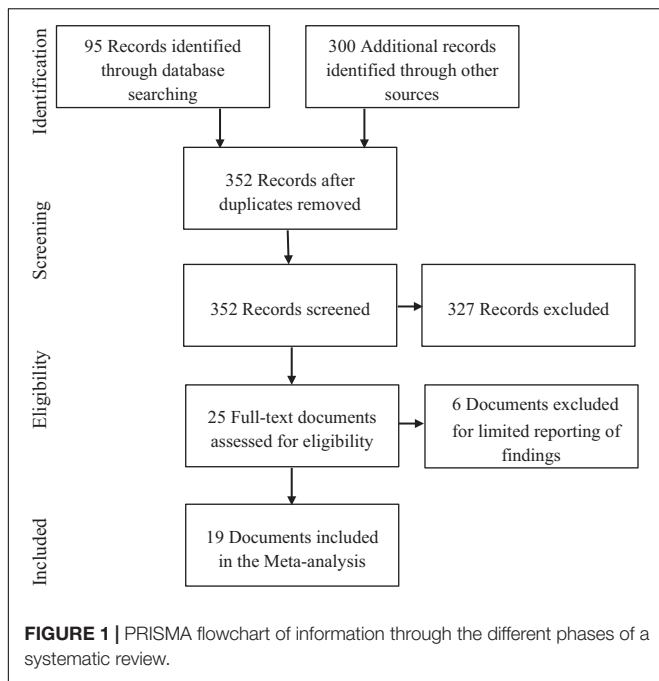
The incremental validity findings concerning the training proficiency are less robust than the findings for job performance. Salgado et al. (2019) showed that CR added validity over CI in the prediction of training proficiency, although the incremental validity was also small ($\Delta R^2 = 0.01$). Meanwhile, Otero (2019) found that CR only added validity over CI in the prediction of GPA ($\Delta R^2 = 0.01$), but not other academic outcomes such as high school grades and college admission scores. Toplak et al. (2014) reported that CR did not add validity over CI in the prediction of GPA. Hence, the fifth research question is:

Research Question 5: Does CR add validity over CI for predicting job performance and training proficiency?

METHOD

Literature Search

We used several strategies to review the literature. The search covered the studies carried out from September 2005 to December 2019. First, we conducted electronic searches in the ERIC database and in the Google Scholar meta-database using the keywords “Cognitive Reflection” and “Cognitive Reflection Test.” Second, we carried out an article-by-article search in the *Journals of Applied Cognitive Psychology, Cognition, Cognitive Science, Frontiers in Psychology, Journal of Applied Research, Journal of Behavioral Decision Making, Journal of Economic Behavior and Operation, Journal of Experimental Psychology: General, Journal of Operations Management, Judgment and Decision Making, Memory and Cognition, Mind and Society, Production and Operations Management, The Journal of Economic Perspectives,* and *The Journal of Socio-Economics, Journal of Behavioral and Experimental Economics, and Thinking and Reasoning.* Third, we checked the reference lists of the papers to identify articles not covered in the abovementioned strategies. Fourth, we



contacted researchers of the CR articles to obtain new studies or supplementary information.

A total of 95 records were identified through the database searches, and 300 additional records were identified through the other three strategies. The content of these 395 records was examined, and the studies that did not provide correlations or data to calculate the correlations between CR and job performance and training proficiency were discarded. The final database consisted of 19 documents. A PRISMA flowchart of information through the different phases of this systematic review appears in **Figure 1**.

Inclusion Criteria and Decision Rules

The content of each record was examined to determine their inclusion in the meta-analysis. We included the studies that provided a correlation or other data that allowed us to estimate the relationship between CR and job performance and training proficiency.

Regarding job performance, we included the correlations between CR and several measures of this occupational criterion. For instance, we included experimental tasks as a proxy of job performance (see Morsanyi et al., 2014; Corgnet et al., 2015b) and the overall performance ratings (see Otero, 2019; Salgado et al., 2019).

For training proficiency, we included the measures of GPA and exam grades. In both cases, the official and the self-reported scores were included. Moreover, knowledge tests required for admission to specific courses were also included as training proficiency measures.

We also included studies where CR was measured with the 3-item CRT or with other CRTs. As other CRTs, we included the tests composed entirely of new items and those which include some of Frederick’s original items, and the CRTs composed entirely of verbal items or combined with numerical ones. The

CRTs composed of three parallel items of the 3-item CRT were not considered as other CRTs, given that the difference between those resides in the nouns that involved (e.g., the bat and ball are substituted for TV and CD; see, for instance, Mata et al., 2013).

When the studies reported, from the same sample, effect sizes administering the 3-item CRT and other CRTs, we included the effect size obtained from 3-item CRT. Then, we also examined the relationship between CR and both organizational criteria according to the CRT type (i.e., 3-item CRT or other CRTs). A meta-analysis using only verbal CRTs could not be conducted because we found hardly any studies.

Based on the above, the meta-analysis on the relationship between CR and job performance was conducted with an accumulated sample size of 1,508 subjects and 7 effects sizes, and the meta-analysis on the relationship between CR and training proficiency was carried out with an accumulated sample of 3,705 subjects and 15 effects sizes. In accordance with MARS and PRISMA guidelines, the primary studies included in the meta-analyses and the relevant information about them (i.e., sample size, observed correlation, predictor reliability, criterion reliability, and the type of the CR test) can be found in **Tables 1, 2** for job performance and training proficiency, respectively.

Meta-Analytic Method

The Schmidt and Hunter’s (2015) psychometric meta-analytic method of correlations was used to analyze the data. This method estimates the true score correlation and the operational correlation between predictor and criterion variables, correcting the mean observed correlations for the artifactual effects. It also estimates the proportion of observed variance (in findings across studies) that is due to the artifacts. The artifacts controlled for in the current meta-analysis were sampling error, measurement error in the predictor variable (i.e., CR), and measurement error in the criterion variable (i.e., job performance and training proficiency). However, because the primary studies rarely provided all the information required to individually correct the observed correlation, it was necessary to create empirical distributions for each artifact.

TABLE 1 | Studies included in the meta-analysis of CR and job performance.

References	Type of CRT	N	r _{xy}	r _{xx}	r _{yy}
^a Čavojsková and Jurkovič, 2017	3-items	164	0.24	0.65	0.80
	7-items	164	0.34	0.76	0.80
Corgnet et al., 2015b	3-items	264	0.34	–	–
Corgnet et al., 2019	3-items	224	0.31	–	–
^a Lado et al., 2021	3-items	245	0.01	0.71	0.90
	13-items	245	0.05	0.79	0.90
^a Otero, 2019	3-items	475	0.30	0.58	0.96
	13-items	475	0.34	0.76	0.96
Salgado et al., 2019	13-items	100	0.33	0.82	0.84
Thoma et al., 2015	7-items	36	0.15	–	–

N, sample size; r_{xy}, observed correlation; r_{xx}, internal consistency reliability of cognitive reflection measure; r_{yy}, internal consistency reliability of job performance.

^aThe observed correlations are provided from the same sample but using different CR test.

TABLE 2 | Studies included in the meta-analysis of CR and training proficiency.

References	Type of CRT	N	r_{xy}	r_{xx}	Study
Corgnet et al., 2015b	3-items	264	0.04	–	
Gómez-Chacón et al., 2014	5-items	56	0.25	–	
Gómez-Veiga et al., 2018	3-items	51	0.32	–	
Insler et al., 2015	3-items	364	0.31	–	
Lindeman and Svedholm-Häkkinen, 2016	3-items	258	0.34	–	
Lohse, 2016	3-items	284	0.27	–	
Morsanyi et al., 2014	3-items	328	0.45	0.57	Study 1
	3-items	184	0.29	0.68	Study 2
Muñoz-Murillo et al., 2020	3-items	189	0.25	–	
^a Otero, 2019	3-items	898	0.16	0.58	Study 3
	13-items	898	0.15	0.78	Study 3
^a Primi et al., 2017	3-items	124	0.17	–	
	6-items	124	0.22	0.76	
Stuppel et al., 2017	3-items	133	0.26	–	
^a Tamas, 2019	3-items	269	0.61	–	
	10-items	269	0.52	–	
^a Thomson and Oppenheimer, 2016	3-items	143	0.33	0.62	
	7-items	143	0.37	0.71	
^a Toplak et al., 2014	3-items	160	0.23	–	
	7-items	160	0.25	0.72	

N, sample size; r_{xy} , observed correlation; r_{xx} , internal consistence reliability of cognitive reflection measure.

^aThe observed correlations are provided from the same sample but using different CR test.

We developed reliability distributions for predictor (i.e., CR) and criterion (i.e., job performance and training proficiency) variables. The distributions were based upon internal consistency coefficients reported in the primary studies (Salgado et al., 2016). In the case of training proficiency, we used the reliability distribution developed by Cuadrado et al. (2021). This distribution was created by 7 coefficients of internal consistency (from 0.78 to 0.98), and the obtained average reliability coefficient was 0.87 ($SD = 0.06$). The mean and the standard deviation of the reliability distributions appear in **Table 3**.

Finally, we corrected the observed mean correlation for measurement error in the criterion variable to obtain the operational correlation (which is relevant for personnel and

TABLE 3 | Predictor and criteria reliability distributions.

	<i>K</i>	r_{xx}	<i>SD</i>	Min.–Max.
Predictor reliability				
CR–job performance	4	0.69	0.10	0.58–0.82
CR–training proficiency	5	0.64	0.08	0.58–0.76
Criteria reliability				
Job performance	4	0.88	0.07	0.80–0.96
Training proficiency ^a	7	0.87	0.06	0.78–0.98

K, number of cases; r_{xx} , average internal consistency reliability; *SD*, the standard deviation of r_{xx} ; Min.–Max., minimum and maximum value of r_{xx} ; CR, cognitive reflection.

^aReliability distribution developed by Cuadrado et al. (2021).

academic decisions), and we corrected the operational correlation for measurement error in the predictor variable to obtain the true score correlation (which is relevant for testing theoretical relationships). The analyses were carried out by Schmidt and Le’s statistical software (Schmidt and Le, 2014).

RESULTS

Meta-Analyses of the CR Validity for Predicting Job Performance and Training Proficiency

Table 4 presents the results of the meta-analysis of the CR validity for predicting job performance. The table, from left to right, presents the number of validity coefficients (*K*), the total sample size (*N*), the observed average correlation (*r*), the observed variance of *r* (S^2_r), the standard deviation of *r* (SD_r), the operational validity (r_{op}), the true validity (ρ), the standard deviation of ρ (SD_ρ), the 90% credibility value (90% CV), the 95% confidence interval of ρ (95% CI), and the percentage of variance explained by artifacts (%VE).

The meta-analytical results show that the observed validity, the operational validity, and the true correlation of CR are 0.28, 0.31, and 0.36, respectively. The values indicate that the CR is a valid predictor of job performance and that the magnitude of the validity is of medium size. Sampling error, CR reliability, and job performance reliability explain 38% of the variance, which suggests that a search for a moderator is in order. The 90% credibility value is notably different from zero, which provides evidence of validity generalization.

Therefore, this first meta-analysis permits to conclude that CR predicts job performance efficiently and that validity generalizes across samples, CR measures, and job performance measures. However, the amount of explained variance is small (13%).

A potential moderator of the CR validity is the type of CRT. In the data set, we observed that studies could be classified into two main groups, a first group of studies that uses the 3-item CRT and a second group of studies that uses other CRTs. Based on this classificatory scheme, we conducted new meta-analyses for the combinations of the type of CRTs and job performance.

Regarding the effects of the type of CRT on validity magnitude and variability, we found that although the observed validity and the operational validity of the 3-item CRT are slightly lower than the respective values for the studies that used CRTs with a higher number of items (0.25 and 0.27 vs. 0.26 and 0.29), the true correlation is slightly higher for 3-item CRT studies than for the studies that used CRT measures with more items (0.33 vs. 0.32). These differences have no practical relevance, and they indicate that the measurement error in the 3-item CRT is higher than that in the other CRTs. Also, the standard deviations and the percentage of explained variance are practically the same for the two groups of studies. Consequently, these results suggest that the type of CRT does not moderate the CR validity for predicting job performance.

TABLE 4 | Meta-analytic results of the CR validity for predicting job performance.

	Bare-bones meta-analysis					Psychometric meta-analysis					
	<i>K</i>	<i>N</i>	<i>r</i>	S^2_r	SD_r	r_{op}	ρ	SD_ρ	90% CV	95% CI	%VE
CR–job performance	7	1,508	0.28	0.0119	0.106	0.31	0.36	0.106	0.22	0.26–0.46	38
CR 3-item	5	1,372	0.25	0.0129	0.114	0.27	0.33	0.129	0.17	0.20–0.47	26
Other CR tests	5	1,020	0.26	0.0154	0.124	0.29	0.32	0.128	0.16	0.19–0.46	28

K, the number of validity coefficients; *N*, the total sample size; *r*, the observed average validity weighted by the study sample size; S^2_r , the observed variance of *r*; SD_r , the standard deviation of *r*; r_{op} , the operational validity; ρ , the true validity; SD_ρ , the standard deviation of ρ ; 90% CV, the 90% credibility value; 95% CI, the 95% confidence interval of ρ ; %VE, the percentage of observed variance explained by artifacts.

TABLE 5 | Meta-analytic results of the CR validity for predicting training proficiency.

	Bare-bones meta-analysis					Psychometric meta-analysis					
	<i>K</i>	<i>N</i>	<i>r</i>	S^2_r	SD_r	r_{op}	ρ	SD_ρ	90% CV	95% CI $_\rho$	%VE
CR–training proficiency	15	3,705	0.27	0.0190	0.138	0.29	0.37	0.168	0.16	0.28–0.47	19
CR 3-item	14	3,649	0.27	0.0193	0.139	0.29	0.37	0.170	0.16	0.28–0.47	18
Other CR tests	6	1,650	0.25	0.0185	0.136	0.26	0.31	0.152	0.11	0.17–0.44	18

K, the number of validity coefficients; *N*, the total sample size; *r*, the observed average validity weighted by the study sample size; S^2_r , the observed variance of *r*; SD_r , the standard deviation of *r*; r_{op} , the operational validity; ρ , the true validity; SD_ρ , the standard deviation of ρ ; 90% CV, the 90% credibility value; 95% CI, the 95% confidence interval of ρ ; %VE, the percentage of observed variance explained by artifacts.

Table 5 shows the results of CR validity for predicting the training proficiency. The estimates for the observed validity, operational validity, and true correlation are 0.27, 0.29, and 0.37, respectively. The 90% CV is 0.16, which indicates that CR generalizes the validity across samples, CR measures, and training proficiency measures. Therefore, as happened with job performance, this analysis allows to conclude that CR is also a robust predictor of the training proficiency.

Concerning the observed variability, the three artifactual errors considered in this meta-analysis explain 19% of the observed variance, which suggests that potential moderators can explain the remaining variance.

As in the case of the job performance criterion, we examined whether the type of CRT moderates the validity of CR for predicting the training proficiency. To this regard, we found that although the estimate for the observed validity, operational validity, and true correlation of the 3-item CRT is slightly higher than the respective values for the studies that used CRTs with a higher number of items (0.27, 0.29, and 0.37 vs. 0.25, 0.26, and 0.31, respectively), these differences have no practical relevance, and they indicate that both shorter and longer CR tests are similarly efficient for predicting training proficiency. Also, the standard deviations and the percentage of variance explained by artifacts are very similar for the two groups of studies. Consequently, these results suggest that the type of CRT does not moderate the CR validity for predicting the training proficiency.

Incremental Validity of CR Over CI for Predicting Job Performance and Training Proficiency

Research Question 5 asked whether the CR shows incremental validity over CI for predicting the two criteria examined in this meta-analysis. Consequently, we carried out hierarchical

multiple regression analyses. The first step consisted of estimating the validity, beta, and explained variance of CI for predicting both criteria. In the second step, we entered CR to estimate the multiple *R*, R^2 , respective betas for CI and CR, and the incremental validity of CR over CI. To conduct these analyses, we used a matrix of true correlations following Schmidt and Hunter's (2015) advice.

In this regard, we created two matrices with the correlations between CI, CR, and the two criteria. We used the values found in this meta-analysis as the correlations between CR and job performance ($\rho = 0.36$) and CR and training proficiency ($\rho = 0.37$). We used the correlation found by the meta-analysis of Otero et al. (2021) as an estimate of the relationship between CI and CR ($\rho = 0.53$). Finally, we used the validities reported by Salgado and Moscoso (2019, **Table 12**) for the true correlations between CI and job performance ($\rho = 0.44$) and CI and training proficiency ($\rho = 0.62$). The harmonic average of all sample sizes was estimated to determine the sample size entered in the regression analysis. According to Judge et al. (2007), the harmonic average is the best estimator of the regression analysis sample size when the purpose is to generalize the results to the population. We decided to follow this recommendation, and the harmonic sample size for the analysis was 4,427.

Table 6 reports the results for the prediction of job performance. As can be seen, CR shows a small (2.5%) but statistically significant incremental validity over CI. The validity changes from 0.440 to 0.465, and the beta weights for CI and CR are also statistically significant. Interestingly, when the two variables are entered in the regression equation, the beta for CI diminishes from 0.440 to 0.347 (a decline of 21%), which indicates that CR is a relevant predictor of job performance, although its contribution is small.

Table 7 reports the results for the prediction of training proficiency when CR supplements CI. As can be seen, the incremental validity of CR is minimal (0.002) and practically

TABLE 6 | Incremental validity of the CR over the CI for predicting job performance.

	β	R	R^2	p	ΔR
Step 1					
CI	0.440	0.440	0.194	0.000	
Step 2					
CI	0.347	0.465	0.216	0.000	0.025
CR	0.176				

$N = 4,427$. All β values are significant at $p < 0.001$ value. CI, cognitive intelligence; CR, cognitive reflection; β , standardized regression weights; R , multiple regression; R^2 , squared multiple correlation; ΔR , incremental validity of the CR over and above the CI.

TABLE 7 | Incremental validity of the CR over the CI for predicting training proficiency.

	β	R	R^2	p	ΔR
Step 1					
CI	0.620	0.620	0.384	0.000	
Step 2					
CI	0.589	0.622	0.387	0.000	0.002
CR	0.058				

$N = 4,427$. All β values are significant at $p < 0.001$ value. CI, cognitive intelligence; CR, cognitive reflection; β , standardized regression weights; R , multiple regression; R^2 , squared multiple correlation; ΔR , incremental validity of the CR over and above the CI.

irrelevant. The validity changes from 0.620 to 0.622, although the beta for CR is statistically significant. When the two variables are entered in the regression equation, the beta for CI diminishes from 0.620 to 0.589 (a decline of 5%), which indicates that CR is not a relevant predictor of training proficiency.

As a whole, these incremental validity analyses show that CR is a useful predictor of job performance but not relevant for predicting the training proficiency.

DISCUSSION

This meta-analytic study examined the current empirical evidence on the relationship between CR and job performance and training proficiency. Moreover, the study used meta-analytically developed correlation matrices to determine the incremental validity of CR over the CI for predicting these two criteria. The meta-analysis aimed to answer five research questions: (1) What is the correlation between CR and job performance? (2) What is the correlation between CR and training proficiency? (3) Is the observed variability in the correlation between CR and job performance and training proficiency real or artifactual? (4) Does the type of CR test moderate the relationship between CR and job performance and training proficiency? (5) Does CR add validity over CI for predicting job performance and training proficiency? In examining the findings of the primary studies with meta-analytical methods, the current research made three unique contributions to the understanding of the relationships between

CR, job performance, and training proficiency and the degree of CR validity generalization and incremental validity.

The first unique contribution has been to show that the CR predicts both job performance and training proficiency. The magnitude of the true correlations is similar or even higher than other widely used predictors of these two criteria, such as personality dimensions, emotional intelligence, specific cognitive abilities, work sample tests, assessment centers, situational and judgment tests, biodata, and social network websites (for instance, Barrick and Mount, 1991; Ones et al., 1993, 2012; Salgado, 1997, 2017; Van Rooy and Viswesvaran, 2004; Joseph and Newman, 2010; Judge et al., 2013; Salgado and Tauriz, 2014; Whetzel et al., 2014; Alonso et al., 2015; Salgado and Lado, 2018; Aguado et al., 2019; Herde et al., 2019; Morillo et al., 2019; Ryan and Derous, 2019; Salgado and Moscoso, 2019; García-Izquierdo et al., 2020; Golubovich et al., 2020; Otero et al., 2020).

The second contribution has been to show that the type of CRTs (i.e., shorter and longer) does not significantly moderate the magnitude of the true correlations. Despite the fact that the short CRT has lower reliability and other psychometric weaknesses, the validity, when corrected for measurement error in CR and criterion, is very similar for this test and other CRTs created to overcome these psychometric limitations.

The third unique contribution has been to show that CR contributes over CI to the explanation of the variance of job performance, although the magnitude of the contribution is small. Despite this fact, the incremental validity and the significant beta indicate that CR should be considered independently from CI when the interest is the prediction of job performance. In the case of training proficiency, CR shows practically no incremental validity over CI validity. The results of multiple regression analysis have indicated that CI supplemented by CR explains the same amount of training proficiency variance as CI alone. This finding contrasts with previous results, which suggested that CR could add validity over CI for predicting GPA (e.g., Otero, 2019; Salgado et al., 2019).

The findings of this meta-analysis have theoretical and practical implications. From the theoretical point of view, the results suggest that CR and CI can be conceptualized as related but different cognitive constructs because they contribute independently to the explanation of job performance.

From a practical point of view, the findings also have relevant implications. On the contrary, as CR correlates significantly with job performance, CRTs could be included among the procedures used for personnel selection, particularly when CI tests were not included in the batteries of selection procedures. CRTs are not a substitute for CI tests, because they represent different constructs, but the criterion validity of the CRTs was of sufficient magnitude for them to be used as an alternative test in the cases in which CI measures are not possible. On the contrary, as CRTs correlate with training proficiency, they can also be used for making decisions in an academic context, for instance, CRT could be used in admission processes (e.g., university studies, training courses, and scholarships).

Like all studies, the present one also has some limitations that must be mentioned. The first limitation is that the number of primary studies is too small for conducting more fine-grain moderator analyses. For example, we were not able to examine

whether job complexity, a relevant moderator of CI validity, is also a moderator of CR validity. The scarcity of studies is also relevant in connection with the measurement of job performance and training proficiency. For example, supervisory performance ratings and instructor ratings, probably the most frequently used measures of these two criteria, were missing in the current database. Therefore, we suggest that future studies include these two types of measures.

CONCLUSION

This research provided the first meta-analysis, which examined the relationships between CR and job performance and training success, showing that CR is an excellent predictor of these two organizational criteria. Moreover, this research showed that CR and CI are not empirically redundant because CR adds validity over CI validity for predicting job performance, although not to predict training proficiency. The results also showed that the type of CRT does not significantly moderate the magnitude of the relationship between CR and job performance and training proficiency. Future research should be conducted to expand the

contributions of this study and to clarify some issues, which were not examined here.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This research was partially supported by the Grant PSI2017-87603-P from the Spanish Ministry of Economy and Competitiveness to JS and SM.

REFERENCES

- Aguado, D., Andrés, J. C., García Izquierdo, A. L., and Rodríguez, J. (2019). LinkedIn “Big Four”: job performance validation in the ICT sector. *J. Work. Organ. Psychol.* 35, 53–64. doi: 10.5093/jwop2019a7
- Alonso, P., Moscoso, S., and Cuadrado, D. (2015). Procedimientos de selección de personal en pequeñas y medianas empresas españolas. *J. Work. Organ. Psychol.* 31, 79–89. doi: 10.1016/j.rpto.2015.04.002
- Baghestanian, S., Lugovskyy, V., and Puzello, D. (2015). Traders’ heterogeneity and bubble-crash patterns in experimental asset markets. *J. Econ. Behav. Organ.* 117, 82–101. doi: 10.1016/j.jebo.2015.06.007
- Barrick, M. R., and Mount, M. K. (1991). The big five personality dimensions and job performance: a meta-analysis. *Pers. Psychol.* 44, 1–26. doi: 10.1111/j.1744-6570.1991.tb00688.x
- Bertua, C., Anderson, N., and Salgado, J. F. (2005). The predictive validity of cognitive ability tests: a UK meta-analysis. *J. Occup. Organ. Psychol.* 78, 387–409. doi: 10.1348/096317905x26994
- Campbell, J. P., and Wiernik, B. M. (2015). The modeling and assessment of work performance. *Annu. Rev. Organ. Psychol. Organ. Behav.* 2, 47–74.
- Campbell, J. P., McHenry, J. J., and Wise, L. L. (1990). Modeling job performance in a population of jobs. *Pers. Psychol.* 43, 313–575. doi: 10.1111/j.1744-6570.1990.tb01561.x
- Campitelli, G., and Labollita, M. (2010). Correlations of cognitive reflection with judgments and choices. *Judgm. Decis. Mak.* 5, 182–191.
- Čavojová, V., and Jurkovič, M. (2017). Comparison of experienced vs. novice teachers in cognitive reflection and rationality. *Stud. Psychol.* 59, 100–112. doi: 10.21909/sp.2017.02.733
- Chaiken, S., Liberman, A., and Eagly, A. H. (1989). “Heuristic and systematic information processing within and beyond the persuasion context,” in *Unintended Thought*, eds J. S. Uleman and J. A. Bargh (New York, NY: Guilford Press), 212–252.
- Corgnet, B., Hernán-González, R., Kujal, P., and Porter, D. (2015a). The effect of earned versus house money on price bubble formation in experimental asset markets. *Rev. Financ.* 19, 1455–1488. doi: 10.1093/rof/rfu031
- Corgnet, B., Hernán-González, R., and Mateo, R. (2015b). Cognitive reflection and the diligent worker: an experimental study of millennials. *PLoS One* 10:e0141243. doi: 10.1371/journal.pone.0141243
- Corgnet, B., Martin, L., Nodjjang, P., and Sutan, A. (2019). On the merit of equal pay: performance manipulation and incentive setting. *Eur. Econ. Rev.* 113, 23–45. doi: 10.1016/j.eurocorev.2018.12.006
- Cuadrado, D., Salgado, J. F., and Moscoso, S. (2021). Personality, intelligence, and counterproductive academic behaviors: a meta-analysis. *J. Pers. Soc. Psychol.* 120, 504–537. doi: 10.1037/pspp0000285
- Evans, J. S. B., and Wason, P. C. (1976). Rationalization in a reasoning task. *Brit. J. Psychol.* 67, 479–486. doi: 10.1111/j.2044-8295.1976.tb01536.x
- Finucane, M. L., and Gullion, C. M. (2010). Developing a tool for measuring the decision-making competence of older adults. *Psychol. Aging* 25, 271–288. doi: 10.1037/a0019106
- Frederick, S. (2005). Cognitive reflection and decision making. *J. Econ. Perspect.* 19, 25–42. doi: 10.1257/089533005775196732
- García-Izquierdo, A. L., Lubiano, M. A., and Ramos-Villagrasa, P. J. (2020). Developing biodata for public manager selection purposes: a comparison between fuzzy logic and traditional methods. *J. Work. Organ. Psychol.* 36, 231–242. doi: 10.5093/jwop2020a22
- Georganas, S., Healy, P. J., and Weber, R. A. (2015). On the persistence of strategic sophistication. *J. Econ. Theory* 159, 369–400. doi: 10.1016/j.jet.2015.07.012
- Golubovich, J., Lake, C. J., Anguiano-Carrasco, C., and Seybert, J. (2020). Measuring achievement striving via a situational judgment test: the value of additional context. *J. Work. Organ. Psychol.* 36, 157–167. doi: 10.5093/jwop2020a15
- Gómez-Chacón, I. M., García-Madruga, J. A., Vila, J. Ó, Elosúa, M. R., and Rodríguez, R. (2014). The dual processes hypothesis in mathematics performance: beliefs, cognitive reflection, working memory and reasoning. *Learn. Individ. Differ.* 29, 67–73. doi: 10.1016/j.lindif.2013.10.001
- Gómez-Veiga, I., Vila, J. O., Duque, G., and García-Madruga, J. A. (2018). A new look to a classic issue: reasoning and academic achievement at secondary school. *Front. Psychol.* 9:400.
- Herde, C., Lievens, F., Solberg, E. G., Harbaugh, J. L., Strong, M. H., and Burkholder, G. J. (2019). Situational judgment tests as measures of 21st century skills: evidence across Europe and Latin America. *J. Work. Organ. Psychol.* 35, 65–74. doi: 10.5093/jwop2019a8
- Hunter, J. E., and Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychol. Bull.* 96, 72–98. doi: 10.1037/0033-2909.96.1.72
- Inslar, M., Compton, J., and Schmitt, P. (2015). The investment decisions of young adults under relaxed borrowing constraints. *J. Behav. Exp. Econ.* 64, 106–121. doi: 10.1016/j.socec.2015.07.004
- Joseph, D. L., and Newman, D. A. (2010). Emotional intelligence: an integrative meta-analysis and cascading model. *J. Appl. Psychol.* 95, 54–78. doi: 10.1037/a0017286

- Judge, T. A., Jackson, C. L., Shaw, J. C., Scott, B. A., and Rich, B. L. (2007). Self-efficacy and work-related performance: the integral role of individual differences. *J. Appl. Psychol.* 92, 107–127. doi: 10.1037/0021-9010.92.1.107
- Judge, T. A., Rodell, J. B., Klinger, R. L., Simon, L. S., and Crawford, E. R. (2013). Hierarchical representations of the five-factor model of personality in predicting job performance: integrating three organizing frameworks with two theoretical perspectives. *J. Appl. Psychol.* 98, 875–925. doi: 10.1037/a0033901
- Kahneman, D. (2011). *Thinking, Fast and Slow*. New York, NY: Farrar, Straus, and Giroux.
- Kahneman, D., and Frederick, S. (2002). “Representativeness revisited: attribute substitution in intuitive judgment,” in *Heuristics of Intuitive Judgment: Extensions and Applications*, eds T. Gilovich, D. Griffin, and D. Kahneman (Cambridge: Cambridge University Press), 49–81. doi: 10.1017/cbo9780511808098.004
- Kahneman, D., and Frederick, S. (2005). “A model of heuristic judgment,” in *The Cambridge Handbook of Thinking and Reasoning*, eds K. J. Holyoak and R. G. Morrison (Cambridge: Cambridge University Press), 267–293.
- Kocher, M., Lucks, K. E., and Schindler, D. (2019). Unleashing animal spirits: self-control and overpricing in experimental asset markets. *Rev. Financ. Stud.* 32, 2149–2178. doi: 10.1093/rfs/hhy109
- Kuncel, N. R., Credé, M., and Thomas, L. L. (2005). The validity of self-reported grade point averages, class ranks, and test scores: a meta-analysis and review of the literature. *Rev. Educ. Res.* 75, 63–82. doi: 10.3102/00346543075001063
- Kuncel, N. R., Hezlett, S. H., and Ones, D. S. (2001). A comprehensive meta-analysis of the predictive validity of the graduate record examinations: implications for graduate student selection and performance. *Psychol. Bull.* 127, 162–181. doi: 10.1037/0033-2909.127.1.162
- Lado, M., Otero, I., and Salgado, J. F. (2021). Cognitive reflection, life satisfaction, emotional balance, and job performance. *Psicothema* 33, 118–124.
- Lindeman, M., and Svedholm-Häkkinen, A. M. (2016). Does poor understanding of physical world predict religious and paranormal beliefs? *Appl. Cogn. Psychol.* 30, 736–742. doi: 10.1002/acp.3248
- Lohse, J. (2016). Smart or selfish—when smart guys finish nice. *J. Behav. Exp. Econ.* 64, 28–40. doi: 10.1016/j.socsec.2016.04.002
- Lubian, D., and Untertrifaller, A. (2013). *Cognitive Ability, Stereotypes and Gender Segregation in the Workplace*. Working Paper No. 25/2013. Italy: University of Verona.
- Mata, A., Fiedler, K., Ferreira, M. B., and Almeida, T. (2013). Reasoning about others’ reasoning. *J. Exp. Soc. Psychol.* 49, 486–491.
- Metcalfe, J., and Mischel, W. A. (1999). A hot/cool-system analysis of delay of gratification: dynamics of willpower. *Psychol. Rev.* 106, 3–19. doi: 10.1037/0033-295x.106.1.3
- Morillo, D., Abad, F. J., Kreitchmann, R. S., Leenen, I., Hontangas, P., and Ponsada, V. (2019). The journey from Likert to forced-choice questionnaires: evidence of the invariance of item parameters. *J. Work. Organ. Psychol.* 2, 75–83. doi: 10.5093/jwop2019a11
- Moritz, B. B., Hill, A. V., and Donohue, K. L. (2013). Individual differences in the newsvendor problem: behavior and cognitive reflection. *J. Operations Manag.* 31, 72–85. doi: 10.1016/j.jom.2012.11.006
- Morsanyi, K., Busdraghi, C., and Primi, C. (2014). Mathematical anxiety is linked to reduced cognitive reflection: a potential road from discomfort in the mathematics classroom to susceptibility to biases. *Behav. Brain Funct.* 10:31. doi: 10.1186/1744-9081-10-31
- Muñoz-Murillo, M., Álvarez-Franco, P. B., and Restrepo-Tobón, D. A. (2020). The role of cognitive abilities on financial literacy: new experimental evidence. *J. Behav. Exp. Econ.* 84:101482. doi: 10.1016/j.socsec.2019.101482
- Noussair, C. N., Tucker, S., and Xu, Y. (2016). Futures markets, cognitive ability, and mispricing in experimental asset markets. *J. Econ. Behav. Organ.* 130, 166–179. doi: 10.1016/j.jebo.2016.07.010
- Oechssler, J., Roeder, A., and Schmitz, P. W. (2009). Cognitive abilities and behavioral biases. *J. Econ. Behav. Organ.* 72, 147–152. doi: 10.1016/j.jebo.2009.04.018
- Ones, D. S., Viswesvaran, C., and Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validities: findings and implications for personnel selection and theories of job performance. *J. Appl. Psychol.* 78, 679–703. doi: 10.1037/0021-9010.78.4.679
- Ones, D. S., Viswesvaran, C., and Schmidt, F. L. (2012). Integrity tests predict counterproductive work behaviors and job performance well: comment on van iddekinge, roth, raymark, and odle-dusseau (2012). *J. Appl. Psychol.* 97, 537–542. doi: 10.1037/a0024825
- Otero, I. (2019). *Construct and Criterion Validity of Cognitive Reflection*. Unpublished doctoral dissertation. Spain: University of Santiago de Compostela.
- Otero, I., Cuadrado, D., and Martínez, A. (2020). Convergent and predictive validity of the big five factors assessed with single stimulus and quasi-ipsative questionnaires. *J. Work. Organ. Psychol.* 36, 215–222. doi: 10.5093/jwop2020a17
- Otero, I., Salgado, J. F., and Moscoso, S. (2021). *Cognitive Reflection, Cognitive Intelligence, and Cognitive Abilities: a Meta-analysis*. Manuscript Submitted for Publication. Department of Political Science and Sociology, University of Santiago de Compostela.
- Primi, C., Morsanyi, K., Chiesi, F., Donati, M. A., and Hamilton, J. (2015). The development and testing of a new version of the cognitive reflection test applying item response theory (IRT). *J. Behav. Decis. Mak.* 29, 453–469. doi: 10.1002/bdm.1883
- Primi, C., Morsanyi, K., Donati, M. A., Galli, S., and Chiesi, F. (2017). Measuring probabilistic reasoning: the construction of a new scale applying item response theory. *J. Behav. Decis. Mak.* 30, 933–950. doi: 10.1002/bdm.2011
- Richardson, M., Abraham, C., and Bond, R. (2012). Psychological correlates of university students’ academic performance: a systematic review and meta-analysis. *Psychol. Bull.* 138, 353–387. doi: 10.1037/a0026838
- Ryan, A. M., and Derous, E. (2019). The unrealized potential of technology in selection assessment. *J. Work. Organ. Psychol.* 35, 85–92. doi: 10.5093/jwop2019a10
- Salgado, J. F. (1997). The five factor model of personality and job performance in the European community. *J. Appl. Psychol.* 82, 30–43. doi: 10.1037/0021-9010.82.1.30
- Salgado, J. F. (2017). Moderator effects of job complexity on the validity of forced-choice personality inventories for predicting job performance. *J. Work Organ. Psychol.* 33, 229–238. doi: 10.1016/j.rpto.2017.07.001
- Salgado, J. F., and Lado, M. (2018). Faking resistance of a quasi-ipsative forced-choice personality inventory without algebraic dependence. *J. Work Organ. Psychol.* 34, 213–216. doi: 10.5093/jwop2018a23
- Salgado, J. F., and Moscoso, S. (2019). Meta-analysis of the validity of general mental ability for five performance criteria: Hunter and Hunter (1984) revisited. *Front. Psychol.* 10:2227.
- Salgado, J. F., Moscoso, S., and Anderson, N. (2016). Corrections for criterion reliability in validity generalization: the consistency of Hermes, the utility of Midas. *J. Work. Organ. Psychol.* 32, 17–23. doi: 10.1016/j.rpto.2015.12.001
- Salgado, J. F., and Tauriz, G. (2014). The five-factor model, forced-choice personality inventories and performance: a comprehensive meta-analysis of academic and occupational validity studies. *Eur. J. Work Organ. Psychol.* 23, 3–30. doi: 10.1080/1359432x.2012.716198
- Salgado, J. F., Anderson, N., Moscoso, S., Bertua, C., de Fruyt, F., and Rolland, J. P. (2003). A meta-analytic study of general mental ability validity for different occupations in the European community. *J. Appl. Psychol.* 88, 1068–1081. doi: 10.1037/0021-9010.88.6.1068
- Salgado, J. F., Otero, I., and Moscoso, S. (2019). Cognitive reflection and general mental ability as predictors of job performance. *Sustainability* 11:6498. doi: 10.3390/su11226498
- Schmidt, F. L., and Hunter, J. E. (2015). *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*, 3rd Edn. Thousand Oaks: Sage.
- Schmidt, F. L., and Le, H. (2014). *Software for the Hunter-Schmidt Meta-Analysis Methods*. Version 2.0. Iowa City, IA: University of Iowa.
- Sirota, M., Kostovičová, L., Juanchich, M., Dewberry, C., and Marshall, A. C. (2020). Measuring cognitive reflection without maths: developing and validating the verbal cognitive reflection test. *J. Behav. Decis. Mak.* 34, 1–22. doi: 10.1002/bdm.2213
- Skagerlund, K., Lind, T., Strömbäck, C., Tinghög, G., and Västfjäll, D. (2018). Financial literacy and the role of numeracy—how individuals’ attitude and affinity with numbers influence financial literacy. *J. Behav. Exp. Econ.* 74, 18–25. doi: 10.1016/j.socsec.2018.03.004
- Stanovich, K. E., and West, R. F. (2002). Variation in how we cope with uncertainty. *Am. J. Psychol.* 115, 127–132. doi: 10.2307/1423679
- Strack, F., and Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Pers. Soc. Psychol. Rev.* 8, 220–247. doi: 10.1207/s15327957pspr0803_1

- Stuppel, E. J., Maratos, F. A., Elander, J., Hunt, T. E., Cheung, K. Y., and Aubeeluck, A. V. (2017). Development of the critical thinking toolkit (CriTT): a measure of student attitudes and beliefs about critical thinking. *Think. Skills Creat.* 23, 91–100. doi: 10.1016/j.tsc.2016.11.007
- Tamas, A. (2019). Correlation investigation: the cognitive reflection test and the math national evaluation scores. *Int. J. Teach. Educ.* 7, 92–103.
- Thoma, V., White, E., Panigrahi, A., Strowger, V., and Anderson, I. (2015). Good thinking or gut feeling? cognitive reflection and intuition in traders, bankers and financial non-experts. *PLoS One* 10:e0123202. doi: 10.1371/journal.pone.0123202
- Thomson, K. S., and Oppenheimer, D. M. (2016). Investigating an alternate form of the cognitive reflection test. *Judgm. Decis. Mak.* 11, 99–113.
- Toplak, M. E., West, R. F., and Stanovich, K. E. (2011). The cognitive reflection test as a predictor of performance on heuristics-and-biases tasks. *Mem. Cogn.* 39, 1275–1289. doi: 10.3758/s13421-011-0104-1
- Toplak, M. E., West, R. F., and Stanovich, K. E. (2014). Assessing miserly information processing: an expansion of the cognitive reflection test. *Think. Reason.* 20, 147–168. doi: 10.1080/13546783.2013.844729
- Van Rooy, D. L., and Viswesvaran, C. (2004). Emotional intelligence: a meta-analytic investigation of predictive validity and nomological net. *J. Vocat. Behav.* 65, 71–95. doi: 10.1016/s0001-8791(03)00076-9
- Wason, P. C., and Evans, J. S. B. (1975). Dual processes in reasoning? *Cognition* 3, 141–154.
- Whetzel, D. L., Rotenberry, P. F., and McDaniel, M. A. (2014). In-basket validity: a systematic review. *Int. J. Select. Assess.* 22, 62–79. doi: 10.1111/ijsa.12057
- Yilmaz, O., and Sarıbay, S. A. (2016). An attempt to clarify the link between cognitive style and political ideology: a non-western replication and extension. *Judgm. Decis. Mak.* 11, 287–300.
- Yilmaz, O., and Sarıbay, S. A. (2017). The relationship between cognitive style and political orientation depends on the measures used. *Judgm. Decis. Mak.* 12, 140–147.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Otero, Salgado and Moscoso. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.