COGNITIVE ABILITIES, ANALYTIC COGNITIVE STYLE AND OVERCONFIDENCE: A COMMENTARY ON DUTTLE (2015)

Michał Białek*+ and Artur Domurat*+

* Kozminski University, Centre for Economic Psychology and Decision Sciences, Jagiellońska 59, 03-301 Warsaw, Poland

+ Both authors share first authorship, they are listed in alphabetical order

Corresponding author:

Michał Białek, Ph.D.

mbialek@kozminski.edu.pl

Kozminski University, Centre for Economic Psychology and Decision Sciences

Jagiellońska 59, 03-301 Warsaw, Poland

ABSTRACT

In his recent paper, Duttle (2015) showed that individuals with higher cognitive abilities show less overconfidence. In these findings, cognitive abilities were equated with an analytic cognitive style (as measured by a cognitive reflection test, or CRT), although recent works in the field of cognitive psychology suggest separating these two constructs. In particular, it is argued that the analytic cognitive style, but not cognitive abilities, decreases susceptibility to cognitive biases. Analyses of data from Duttle's study support this assertion. Implications for cognitive psychology and behavioral economics are discussed.

Cognitive abilities, analytic cognitive style and overconfidence: A commentary on Duttle (2015)

1. Introduction

Psychological science has investigated cognitive bias and its origins at least since the heuristics and biases paradigm was established in the mid-70s by Amos Tversky and future Nobel Prize winner Daniel Kahneman. The basic finding is, that biases are universal, automatic, and only weakly affected by effort or explicit instructions (Stanovich & West, 2008). Thus, they are "mounted" or "hardwired" in the human mind (De Neys, 2006; Evans & Frankish, 2009). Contrary to this view, in his recent paper, Duttle (2015) reported a study of cognitive abilities' (CA) impact on overconfidence, endorsing the idea that higher CA promote better calibration of confidence. This claim is thought provoking and theoretically intriguing. Overconfidence is considered to be a precondition of other biases, even having been called "a mother of all biases" (Bazerman & Moore, 2013). For this reason, our reading of the paper leads us to disagree with the conclusions drawn by Duttle.

In the aforementioned work, the cognitive reflection test (CRT, Frederick, 2005) was intended to measure CA. Two methods served as measures of overconfidence: self-estimated performance in Raven's Progressive Matrices (RPM; Raven, 2000) and a set of interval estimates adapted from Glaser, Langer, and Weber (2013). Duttle investigated three types of overconfidence: overplacement (thinking of oneself as better than average), overestimation (estimating one's own performance as better than it really is) and overprecision (being too certain that one's beliefs are accurate and correct). The main finding was that CRT score was associated with a significant decrease in biases such as overplacement and overprecision, but not overestimation. The author concluded that CA promote better calibration of confidence.

In our opinion, Duttle's paper has important limitations. Specifically, we address three concerns: (1) a confound between CA and analytic cognitive style (ACS); (2) the use of differences between true and subjectively perceived performance in the RPM as a measure of overconfidence; and, in consequence, (3) misinterpretation of reported statistical analysis. We shall show that the relationship between CA and overconfidence disappears after reinterpreting the analysis provided by Duttle.

2. The CRT does not measure CA

A better understanding of the problems associated with using the CRT as a measure of CA can be obtained by considering the dual-process theory of thinking. This posits that the human mind uses two types of processes. Type 1 processes are the default type; they are automatically triggered, require no cognitive resources, and are fast and intuitive. Although they cannot be omitted, they can be overridden by type 2 processes, which require working memory resources and are slow and rule-based (Evans & Stanovich, 2013; Kahneman, 2011).

The difference between CA and the ACS is apparent in the relationship between these two systems. CA relate to the efficiency of type 2 processes, while ACS positioning determines the willingness to engage in type 2 thought processes.

Since the development of the CRT, it has come to be excessively used, with studies showing that it correlates with a plethora of behavioral and cognitive factors. For example, it correlates with such distal constructs as religious disbelief, paranormal beliefs and moral judgments (for reviews, see Pennycook, Fugelsang, & Koehler, 2015a). In this test, individuals solve simple mathematical tasks, which may tempt researchers to use it as a measure of mathematical abilities or of cognitive abilities in general. This, however, is not a valid approach (Pennycook & Ross, 2016). In fact, CRT measures the willingness to engage in deliberation, and such deliberation results in increased performance in cognitive ability tests like RPM. Thus, although previous research shows that CRT scores correlate with verbal intelligence, general intelligence and reasoning performance with a strength of .20 – .50 (see, e.g., Pennycook, Ross, Koehler, & Fugelsang, 2016; Toplak, West, & Stanovich, 2011, 2014; Trippas, Pennycook, Verde, & Handley, 2015), these variables should not be unified.

In short, cognitive reflection moderates the impact of cognitive abilities on related tests (Trippas et al., 2015). Thus, it is unsuitable for testing the impact of CA on overconfidence. On one hand, high ACS reduces susceptibility to cognitive biases (Noori, 2016; Stanovich, 2009); in our opinion, this is a reason why Duttle registered a negative relationship between CRT and overconfidence. On the other hand, we will explain below that cognitive abilities do not prevent biases. Even Duttle himself admits that CRT is used as a measure of analytic cognitive style rather than of cognitive abilities; however, he concludes that *"the CRT is utilized in a number of studies that investigate cognitive skills [...] since despite its simplicity CRT scores quite accurately predict cognitive abilities, in particular with respect to mathematical skills. As a matter of simplicity, in the remainder of the text the CRT therefore is referred to as a measure of cognitive skills*" (Duttle, 2015, p. 5). Although a similar approach has been taken in some other economical works (Oechssler, Roider & Schmitz, 2009; Hoppe & Kusterer, 2011; Brañas-Garza, Garcia-Muñoz &, González, R. H., 2012), we cannot find any justification for substituting one cognitive notion with another psychologically distinct concept (especially if these measures correlate merely at r = .4, as reported in the aforementioned work). Proposing that cognitive abilities predict less overconfidence is conceptually different from stating that cognitive style predicts it.

Fortunately, Duttle used a theoretically well-grounded measure of CA, which is the RPM. In our opinion, this is a major but inefficiently exploited advantage of Duttle's study, and we will return to this issue when reinterpreting and remodeling his findings.

3. Cognitive abilities do not prevent biases

The important implication of dual process theory is that, despite their ability to perform numerous cognitive operations, some people fail to engage in deliberation and instead rely on intuitive responses. Cognitive biases are produced predominantly when type 1 processes are not overridden by type 2 processes (Stanovich, Toplak, & West, 2008; Toplak, West, & Stanovich, 2011). This is why individuals with high CA show the same vulnerability to cognitive biases as individuals with low CA (De Neys & Glumicic, 2008; Evans, 2007; Pennycook, Fugelsang, & Koehler, 2015b). Thus, an analytic cognitive style, rather than CA, predicts reduced susceptibility to cognitive biases (Noori, 2016; Toplak et al., 2011), and the positive impact of CA on some tasks (e.g., reasoning) can be moderated by ACS (Trippas, et al. 2015). Specifically, considering the cited literature, CA should have a positive impact on thinking only when people are willing to think analytically.

As we explained above, by showing that CRT scores impact overconfidence, Duttle drew an unjustified conclusion that cognitive abilities reduce overconfidence. However, having been provided with valid measures of cognitive abilities (RPM), cognitive style (CRT) and overconfidence, we can remodel the findings and verify his hypothesis stating that cognitive abilities prevent overconfidence. Prior to such modeling, we need to discuss the validity of dependent variables, namely overplacement. We think that its design artificially creates a correlation between overconfidence and cognitive abilities, as both are based on performance in the RPM.

4. True scores on Raven's matrices shape overconfidence measures in RPM performance

We agree with Duttle that tasks such as intelligence tests are a great way to capture overconfidence, as "*the effect of cognitive abilities to be most informative in an environment where they actually are an important factor of success*" (Duttle, 2015, p. 12). Individuals are more likely to have a well-calibrated self-image in the context of such important factors, as compared to self-image in less important tasks (e.g., those measuring general knowledge unrelated to everyday problems). We also agree that "*Compared to methods implementing general knowledge questions and other measures of overconfidence that are commonly used in laboratory studies [such method] is completely independent from participants' education and training*" (p. 3), using methods related to important everyday life activities created better conditions to test the relationship between the performance and overconfidence.

Having the aforementioned advantages, usage of RPM to capture overconfidence has also a methodological concern. Specifically, a score on a test and – as a measure of overconfidence – the difference between a self-assessment in this test and the score are artificially correlated. In the cited study, participants estimated their RPM performance on a scale ranging from 0 to 10. The difference between the estimation and the actual performance served as a measure of overestimation; thus, potential scores could range from -10 to +10, with lower and higher scores indicating underestimation and overestimation, respectively. An analogical measure was constructed for overplacement; the difference between the estimated quintile and the actual quintile scored potentially from -4 to +4.

Duttle reported a strong negative correlation between RPM score and the RPM-based overestimation index, $r(97) = -.69$. We suggest that the scoring of RPM artificially shapes the range of possible overconfidence scores. RPM is an intelligence evaluation tool. Thus, individuals with higher and lower CA respectively score higher and lower on RPM, respectively, and consequently have respectively less and more space to manifest overconfidence. Consider the example of two fictitious people: Mario, who scores 9 and is placed in the highest RPM quintile, and Luigi, who scores 2 and is placed in the lowest quintile. Both estimate their performance on a 0-10 scale for the number of correct RPM items and 1-5 for the quintile comparing their result to other participants. Mario's overestimation score can thus range from -9 to +1, while Luigi's can range from -2 to +8. The same applies to overplacement, where Mario can score from -4 to 0 and Luigi from 0 to +4. These boundaries create substantial artificial negative correlations between RPM-based overconfidence measures and RPM scores, which would occur even if people estimated their RPM performance and placement at random. Such a correlation is not essential for majority of studies, in which overconfidence is measured on tasks like general knowledge quizzes. Even if the performance on such a quiz affects the relative overconfidence, the relationship between overconfidence and other variables remains unaffected. Stankov and Crawford (1997) reported a relevant study that investigated overconfidence measured on RPM with the performance on RPM. Using signal detection theory and receiver operating characteristics rather than linear regression, the authors concluded that individual showed almost no overconfidence.

In sum, in this particular study, the RPM, which served as a basis for estimating overconfidence, is also a measure of CA (proposed predictor of the overconfidence). Considering their artificial correlation, we cannot validly interpret the regressions reported by Duttle. We can, however, simulate what would happen assuming a particular relationship between both the variables and constraints of the spectrum of responses as in the Duttle study and comparing it to the correlation obtained.

4.1. Testing the impact of RPM-based cognitive abilities on the RPM-based overconfidence

Having in mind the artificial negative correlation between performance in the RPM and overconfidence, we intended to compare the reported correlation with a dataset from simulated experiments in which the relationship between variables is explicitly specified. Specifically, we simulated sets of 1000 experiments, each with 100 random observations (to obtain the same significance criterion for Pearson's *r* coefficients as in Duttle's study), with each set having different base assumptions (see Table 1). Additionally, a control condition of perfect self-assessment is included, in which the correlation is by definition equal to zero (no overconfidence).

Table 1. Simulated average correlation coefficients between overestimation and Ravens progressive matrices score and their difference to the coefficient reported in Duttle (2015), r = -.69.

A sample of:	Condition	Average simulated correlation between RPM quintile and overplacement measure	Fishers z-test, comparing Duttle's result with the typical coefficient
Ideal respondents	Individuals perfectly estimate their real performance	r = 0	5.91, p < .001
Motivated respondents	50% of individuals commit an error by a maximum of 1 quintile, 50% of individuals have well-calibrated self-evaluations	r = - 0.26 (-.46 to .01)	4.05, p < .001
Guessing respondents	Assume totally random self-evaluations	r = - 0.71 (-.81 to -.54)	-0.27, p = .787

From Table 1 we can see that, regardless of assumed ability to estimate one's own performance, the relationship between RPM-based overestimation and true RPM score in all simulations is non-positive. Because of the same theoretical issues, similar artificial correlation occurs for overplacement scores. Moreover, as showed by simulation, the correlation coefficient reported in Duttle's study, r = -.69, is significantly different from all other samples but from the random-responding individuals, thus suggesting that, in reality, RPM score is probably only very weakly correlated, or even not correlated at all, with the overestimation and overplacement of one's own performance in this task; thus, as mentioned previously, the regression analysis conducted by Duttle reports an artificial relationship between cognitive abilities measured by the RPM and overconfidence.

4.2. Reinvestigating the relationship between cognitive abilities and overplacement and overprecision

The arguments presented in section 2 suggest that CRT scores' relationship with overconfidence cannot be interpreted in the manner done by Duttle (2015), namely that this relationship evidences the impact of cognitive abilities on overconfidence. However, Duttle provided enough data to test his hypotheses, which state that RPM scores are a valid measure of CA and that the interval estimation task provides a valid measure of overconfidence (i.e., overplacement and overprecision). Interestingly, Duttle conducted an analysis that directly tests the effects of cognitive abilities on overconfidence with control of the analytic cognitive styles, yet it is presented merely as a supplement to the main text. Discussing the results of a restricted tobit regression (where RPM and CRT scores are regressed on overprecision and overplacement), he concludes that:

> "*We can observe that the CRT score is a much better predictor of both overprecision and overplacement than performance in the RPM. An impulsive reasoning style appears to be correlated with high overprecision in interval forecasts and with low comparative overplacement. Interestingly, when controlling for CRT scores, general deductive ability measured in the RPM does not explain overconfidence at all" (p. 3, Appendix).*

Again, keeping in mind that CRT measures ACS and RPM measures CA, we can clearly see that Duttle's results are consistent with previously reported experiments and do not support the author's main conclusion. As an aside, it is worth noting that CRT cannot be used as a measure of "impulsive reasoning style", but rather should be used to measure "reflective reasoning style", as intuitive errors in the CRT have no predictive value (Pennycook, Cheyne, Koehler, & Fugelsang, 2015). Thus, it should be stated instead that reflective cognitive style appears to be negatively correlated with overprecision in interval forecasts and with comparative overplacement. Such findings have recently been reported in the studies of Oechssler, et al. (2009), Hoppe and Klusterer (2011), and Noori (2016), who showed that less reflective individuals tend to be more overconfident and have less calibrated self-assessments in tests that are orthogonal to cognitive abilities.

5. Conclusions

Our rethinking of Duttle's (2015) experiment suggests that there is no substantial evidence that cognitive abilities have an impact on overconfidence. Duttle's interpretation would concur with the current psychological literature proposing that cognitive biases arise from the general architecture of the human mind rather than low cognitive ability (De Neys, 2006; Stanovich & West, 2008). Specifically, individuals usually rely on intuitive cognitive processing and heuristic thinking. However, under certain conditions, they can engage in more effortful processing and override intuition with effortful, rule-based thinking (Kahneman, 2011; Pennycook et al., 2015b; Handley & Trippas, 2015). This is referred to as the default interventionist model of cognition and enjoys current broad acceptance among cognitive scientists (Evans, 2011; Evans & Stanovich, 2013). We conclude that Duttle's study contributes to this approach, as his data show that the ability to overcome cognitive biases stems from a *willingness* to test one's own intuitions (analytic cognitive style) rather than from an *ability* to perform well in complex cognitive tasks (cognitive abilities).

Acknowledgments

We would like to thank Kai Duttle for providing his dataset and comprehensive explanations of his computations, and the reviewers for great deal of helpful comments and remarks.

References

Bazerman, M. H., & Moore, D. A. (2013). *Judgment in managerial decision making* (8 ed.). Hoboken: Wiley & Sons.

Brañas-Garza, P., Garcia-Muñoz, T., González, R. H. (2012) Cognitive effort in the beauty contest game. *Journal of Economic Behavior & Organization* 83, 254-260.

De Neys, W. (2006). Dual processing in reasoning: two systems but one reasoner. *Psychological Science*, 17, 428-433.

De Neys, W., & Glumicic, T. (2008). Conflict monitoring in dual process theories of thinking. *Cognition*, 106, 1248-1299.

Duttle, K. (2015). Cognitive skills and confidence: interrelations with overestimation, overplacement and overprecision. *Bulletin of Economic Research*, in press.

Evans, J. S. B. (2007). On the resolution of conflict in dual process theories of reasoning. *Thinking & Reasoning*, 13, 321-339.

Evans, J. S. B. (2011). Dual-process theories of reasoning: Contemporary issues and developmental applications. *Developmental Review*, 31, 86-102.

Evans, J. S. B., & Frankish, K. E. (2009). *In two minds: Dual processes and beyond*. Oxford: Oxford University Press.

Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition advancing the debate. *Perspectives on Psychological Science*, 8, 223-241.

Frederick, S. (2005). Cognitive Reflection and Decision Making. *The Journal of Economic Perspectives*, 19, 25-42.

Glaser, M., Langer, T., & Weber, M. (2013). True overconfidence in interval estimates: Evidence based on a new measure of miscalibration. *Journal of Behavioral Decision Making*, 26, 405-417.

Handley, S. J., & Trippas, D. (2015). Dual Processes and the Interplay between Knowledge and Structure: A New Parallel Processing Model. In B. Ross (Ed.), *Psychology of Learning and Motivation* (pp. 33-58). Burlington: Academic Press

Hoppe E. I., Kusterer D. J. (2011) Behavioral biases and cognitive reflection. *Economics Letters* 110, 97-100.

Kahneman, D. (2011). *Thinking Fast and Slow*. New York: Farrar, Strauss and Giroux.

Noori, M. (2016). Cognitive reflection as a predictor of susceptibility to behavioral anomalies. *Judgment and Decision Making*, 11, 114-120.

Oechssler, J., Roider, A., & Schmitz, P. W. (2009). Cognitive abilities and behavioral biases. *Journal of Economic Behavior & Organization*, 72, 147-152.

Pennycook, G., Cheyne, J. A., Koehler, D. J., & Fugelsang, J. A. (2015). Is the cognitive reflection test a measure of both reflection and intuition? *Behavior Research Methods*, 48, 341-348..

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015a). Everyday consequences of analytic thinking. *Current Directions in Psychological Science*, 24, 425-432.

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015b). What makes us think? A three-stage dual-process model of analytic engagement. *Cognitive Psychology*, 80, 34-72.

Pennycook, G., & Ross, R. M. (2016). Commentary: Cognitive reflection vs. calculation in decision making. *Frontiers in Psychology*, 7, 9.

Pennycook, G., Ross, R. M., Koehler, D. J., & Fugelsang, J. A. (2016). Atheists and Agnostics Are More Reflective than Religious Believers: Four Empirical Studies and a Meta-Analysis. *PloS One*, 11, e0153039.

Raven, J. (2000). The Raven's progressive matrices: change and stability over culture and time. *Cognitive Psychology*, 41, 1-48.

Stankov, L., & Crawford, J. D. (1997). Self-confidence and performance on tests of cognitive abilities. *Intelligence*, 25, 93-109.

Stanovich, K. E. (2009). *What intelligence tests miss: The psychology of rational thought.* Yale University Press.

Stanovich, K. E., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology*, 94, 672-695.

Stanovich, K. E., Toplak, M. E., & West, R. F. (2008). The development of rational thought: A taxonomy of heuristics and biases. *Advances in Child Development and Behavior*, 36, 251.

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, 1275-1289.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking & Reasoning*, 20, 147-168.

Trippas, D., Pennycook, G., Verde, M. F., & Handley, S. J. (2015). Better but still biased: Analytic cognitive style and belief bias. *Thinking & Reasoning*, 21, 431-445.