Developing A Cognitive Reflection Test for School-Age Children

Andrew G. Young (ayoung2@oxy.edu), Allison Powers, Lesley Pilgrim,
& Andrew Shtulman (shtulman@oxy.edu)
Department of Psychology, Occidental College
1600 Campus Road, Los Angeles, CA 90041

Abstract

The cognitive reflection test (CRT; Frederick, 2005) assesses how well adults can reflect on the validity of their own thinking, and it has been shown to predict several measures of normative reasoning. Here, we sought to create a version of the cognitive reflection test suitable for elementary-school-aged children, which could be used to study the emergence of cognitive reflection as well as its role in the development of other forms of higher-order cognition. We identified eight child-friendly questions that elicit an incorrect, intuitive response that must be inhibited in order to provide a correct, analytic response. We compared children's and adults' performance on these questions (dubbed the CRT-D) to several measures of rational thinking (denominator neglect, base rate sensitivity, syllogistic reasoning, otherside thinking) and thinking dispositions (actively open-minded thinking, need for cognition). The CRT-D was a significant predictor of rational thinking and normative thinking dispositions in both children and adults. Moreover, performance on the CRT-D correlated with performance on the original CRT in adults, and in children, it predicted rational thinking and normative thinking dispositions above and beyond age. These results suggest that the CRT-D is a valid measure of children's cognitive reflection and pave the way for future investigations of its development and its developmental consequences.

Keywords: cognitive reflection, rational thinking, cognitive development

Introduction

The cognitive reflection test (CRT; Frederick, 2005) is among the most widely-used measures of adults' tendency to engage in analytic vs. intuitive thinking. The measure was designed to measure a person's tendency to override a gut intuitive response that is incorrect and to engage in reflection that leads to a correct response. Consider the well-known bat and ball item: "A bat and a ball cost $1.10 in total. The bat costs $1 more than ball. How much does the ball cost?" Many adults provide the intuitively elicited response of 10 cents, defaulting to simple subtraction. However, the correct answer is 5 cents (because the bat must cost $1.05 in order for their sum to be $1.10 and their difference $1.00), and adults who provide that answer have presumably engaged in analytical thinking, realizing that the intuitive response that first came to mind was incorrect and generating a correct response in its place.

Performance on the CRT strongly predicts performance on a variety tasks and measures. For example, the CRT correlates with rational thinking on many standard heuristics and biases tasks, such as belief bias in syllogistic reasoning, denominator neglect, otherside thinking, framing, and temporal discounting (Frederick, 2005; Toplak, West, & Stanovich, 2014a). Indeed, Toplak, West, and Stanovich (2011) found a composite variable of 15 different rational thinking tasks was better predicted by the CRT than by either general intelligence or executive functioning. The CRT also correlates with a number of thinking disposition measures, including need for cognition, actively open-minded thinking, and superstitious thinking (e.g., Frederick, 2005; Toplak et al., 2014a). Moreover, the CRT predicts cognitive and behavioral measures from other domains of psychology, such as science understanding, religious and paranormal belief, moral reasoning, creativity, and pro-sociality (see Pennycook, Fugelsang, & Koehler, 2015).

Broad utility and interest in the CRT have led researchers to develop a number of alternative versions. These versions address known problems with the original CRT, such as floor effects in adolescents and certain adult populations, the items' heavy reliance on numeracy, and the general public's increased familiarity with the items themselves (e.g., Baron, Scott, Fincher, & Metz, 2015; Primi, Morsany, Chiesi, Donati, & Hamilton, 2015; Stanovich, West, & Toplak, 2016; Thomson & Oppenheimer, 2016; Toplak et al., 2014a). The goal of the current research was to develop a version of the CRT for school age children. A successful measure of children's cognitive reflection would make important contributions to several lines of research. First, as with adults, an individual difference measure of children's analytic vs. intuitive cognitive style might be used to investigate children's reasoning across several domains, including scientific reasoning, moral reasoning, and social reasoning. Second, such a measure could be used to investigate the developmental trajectory of cognitive reflection and rational thinking, as well as its malleability with experience or instruction (e.g., Kokis, Macpherson, Toplak, West, & Stanovich, 2002; Primi et al., 2015; Toplak, West, & Stanovich, 2014b). Finally, a developmental measure might help adjudicate between competing accounts of cognitive reflection in adults, such as whether producing a counterintuitive response requires first inhibiting the intuitive response or merely reflects the stronger activation of one response over another (Travers, Rolison, & Feeney, 2016).

The present study proposed and validated a developmental version of the cognitive reflection test (CRT-D). We developed a set of 8 cognitive reflection items that were similar in structure to the original CRT. In particular, each problem had an intuitive (but incorrect) lure and a correct response that we expected school age children would be capable of producing upon reflection. For instance, one of our items was "What do cows drink?" The intuitive, lure response is "milk," but the correct response is actually "water." Following recent research on the development of

Table 1. CRT-D questions and percentages of response types (correct; intuitive; other)

Question	Adults	Children
1. If you're running a race and you pass the person in second place, what place are you in? (correct: second; intuitive: first)	72; 27; 0	21; 67; 12
2. Emily's father has three daughters. The first two are named Monday and Tuesday. What is the third daughter's name? (correct: Emily; intuitive: Wednesday)	70; 28; 2	11; 66; 23
3. A farmer has 5 sheep, all but 3 run away. How many are left? (correct: three; intuitive: two)	90; 8; 2	19; 76; 5
4. If there are 3 apples and you take away 2, how many do you have? (correct: two; intuitive: one)	96; 4; 0	46; 46; 8
5. What do cows drink? (correct: water; intuitive: milk)	91; 9; 0	53; 44; 3
6. What weighs more, a pound of rocks or a pound of feathers? (correct: same weight; intuitive: rocks)	88; 12; 0	7; 92; 1
7. What hatches from a butterfly egg? (correct: caterpillar; intuitive: baby butterfly)	72; 19; 9	67; 28; 5
8. Who makes Christmas presents at the North Pole? (correct: elves; intuitive: Santa)	73; 27; 0	41; 58; 1
[Not Included] What do you put in a toaster? (correct: bread; intuitive: toast)	88; 2; 10	58; 28; 14

newer versions of the CRT (Baron et al., 2015; Primi et al., 2015; Stanovich et al., 2016; Thomson & Oppenheimer, 2016; Toplak et al., 2014a), we investigated the validity of the CRT-D by assessing whether the CRT-D was similar to the original CRT in its relationships to other measures. First, a sample of adults completed our CRT-D, the original CRT, a set of items from newer versions of the CRT, and a battery of rational thinking and thinking disposition measures that have previously been found to correlate with the CRT. For adults, the question of interest was whether the new CRT-D items would correlate with these measures similarly to the original CRT. Second, a sample of school-age children completed the CRT-D as well as age appropriate versions of the same rational thinking and normative thinking disposition measures (e.g. Toplak et al, 2014b). For children, the question of interest was whether the CRT-D items would yield correlations with rational thinking and thinking disposition measures that frequently correlate with the CRT in adults.

Method

Participants

One hundred adults participated via Amazon's Mechanical Turk. Ninety-six 5- to 12-year-old children participated (*M* age = 8 years, 1 month, *SD* = 2 years, 2 months, 49% female). The age distribution of child participants is shown in Figure 1. Children were recruited from public playgrounds and completed the study onsite. Eleven children did not complete the full battery of measures due to equipment failure, parent interruption, or attrition. Thus, we report child sample sizes for each measure below.

Procedure and Materials

Participants completed the battery of tasks described below in the following order: CRT-D, denominator neglect, belief bias syllogisms, base rate sensitivity, otherside thinking, need for cognition, and actively open-minded thinking. Adults additionally responded to alternative CRT questions intermixed with the CRT-D and the original CRT questions following the otherside thinking task. Adults completed the study online. Children completed the study one-on-one with trained research assistants. Research assistants read aloud items as they were displayed on an iPad. Children responded verbally or via touch screen depending on the measure.

CRT-D. Participants answered 8 cognitive reflection items for children that were similar in structure to the original CRT. Three items (Questions 1-3) were adapted from the non-numerical CRT developed by Thomson and Oppenheimer (2016). The remaining questions were found by searching for children's "brain teasers" online. An additional item ("What do you put in a toaster?") did not possess the fundamental attribute of the original CRT items: that the vast majority of participants generate either a single correct or single intuitive/heuristic incorrect response. Children and adults generated many correct alternative responses (e.g., bagels and waffles), and thus, we did not include this item in the final 8-item scale. (Child *n* = 96)

CRT and CRT-Alt. Adults answered the original 3-item CRT (Frederick, 2005) as well as 6 items taken from published extended versions of the CRT (Primi et al, 2015; Toplak et al., 2014a; Stanovich et al., 2016).

Denominator Neglect. Participants were shown trays of black and white marbles and told that black marbles were

Table 2. Descriptive statistics of variables and correlations among variables for adults.

Variables	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
1. CRT-D	.	.	.	.	.	.	.	.	.	.	.
2. CRT	.47***	.	.	.	.	.	.	.	.	.	.
3. CRT-Alt	.44***	.53***	.	.	.	.	.	.	.	.	.
4. Denominator neglect	.21*	.17+	.37***	.	.	.	.	.	.	.	.
5. Base rate sensitivity	.23**	.35***	.30***	.25*	.	.	.	.	.	.	.
6. Belief bias syllogisms	.32**	.36***	.50***	.43***	.26**	.	.	.	.	.	.
7. Otherside thinking	-.04	.00	.00	.01	.07	.05	.	.	.	.	.
8. Rational thinking composite	.29**	.30**	.27***	.68***	.64***	.69***	.42**	.	.	.	.
9. Actively Open-Minded Thinking	.04	.19+	.27**	.36***	.22*	.30**	.19+	.35***	.	.	.
10. Need for Cognition	.08	.05	-.03	.15	.09	-.06	.00	.03	.25*	.	.
11. Normative thinking disposition composite	.08	.15	.15	.32**	.20*	.16	.12	.24*	.79***	.79***	.
Mean	6.51	2.34	3.17	8.20	4.08	5.90	4.25	.00	3.15	2.90	.00
SD	1.25	1.03	1.84	1.60	1.06	2.16	2.09	.58	.43	.62	.79
Cronbach's alpha	.31	.76	.72	.81	.52	.78	.	.	.91	.75	.
Split-half reliability	.31	.44	.65	.79	.45	.76	.	.	.89	.75	.

+ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$

the "winners"—i.e., that the goal of the game was to select a black marble. They were then asked to choose between a smaller tray which contained fewer winning marbles but a higher probability of winning, or a larger tray with more winning marbles and a lower probability of winning (e.g., 1:10 vs. 9:100). Participants responded to 9 items of varying ratios. To reduce tendency toward a response set, 3 of these items involved the larger tray also having a higher probability of winning. Instructions were from a developmental version of the task (Toplak et al., 2014b). (Child *n* = 92)

Belief Bias Syllogisms. Participants evaluated the logical validity of 8 syllogisms consisting of either an invalid argument paired with a believable conclusion (e.g., "All flowers have petals; Roses have petals; Roses are flowers") or a valid argument paired with an unbelievable conclusion (e.g., "All mammals walk; Whales are mammals; Whales walk"). Items and instructions were from a developmental version of the task (Toplak et al., 2014b). (Child *n* = 92)

Base Rate Sensitivity. Participants evaluated 5 scenarios in which probabilistic base rate information conflicted with concrete/personal information. An example is "Erica wants to go to a baseball game to try to catch a fly ball. She calls the main office and learns that almost all fly balls have been caught in section 43. Just before she chooses her seats, she learns that her friend Jimmy caught a fly ball last week sitting in section 10. Which section is most likely to give Erica the best chance to catch a fly ball?" Items and instructions were from a developmental version of the task (Kokis et al., 2002; Toplak et al., 2014b). (Child *n* = 91)

Otherside Thinking. Participants completed a standard otherside thinking task with a child-appropriate topic (Toplak et al., 2014b). Participants were asked to give their position on the following issue: "Do you think kids should have cell phones?" Participants were then asked to give reasons for and against their position. The key measure was the number of conceptually unique reasons a participant provided against their endorsed position. (Child *n* = 91)

Need for Cognition. Adults completed a 9-item need for cognition scale (NFC) devised by Kokis et al. (2002). Children completed a NFC scale more recently developed and validated for children and adolescents (Keller et al., 2016). Examples from the 14-item child scale are "Thinking is fun for me" and "I like learning new things." Children and adults responded on a 4-point agreement scale, with higher scores indicating greater motive to engage in effortful cognitive activities. (Child *n* = 87)

Actively Open-Minded Thinking. Adults completed the 7-item actively open-minded thinking scale (AOT) of Haran, Ritov, and Mellers (2013). Children responded to a version in which we modified several items to be more child-friendly. Examples of child-modified items are "It is good to listen to the other side of an argument" and "Changing your mind is a bad thing" (reverse scored). Participants responded on a 4-point agreement scale, with higher scores indicating a greater tendency towards open-minded thinking. (Child *n* = 87)

Table 3. Descriptive statistics of variables and correlations among variables for children.

Variables	1.	2.	3.	4.	5.	6.	7.	8.	9.
1. CRT-D	.	.	.	.	.	.	.	.	.
2. Denominator neglect	.31**	.	.	.	.	.	.	.	.
3. Base rate sensitivity	.37***	.25*	.	.	.	.	.	.	.
4. Belief bias syllogisms	.06	-.06	.12	.	.	.	.	.	.
5. Otherside thinking	.19+	.09	.01	.07	.	.	.	.	.
6. Rational thinking composite	.41***	.59***	.62***	.51***	.53***	.	.	.	.
7. Actively Open-Minded Thinking	.30**	.24*	.24*	-.01	.11	.26*	.	.	.
8. Need for Cognition	.19+	.09	-.12	-.22*	.24*	.00	.11	.	.
9. Normative thinking disposition comp.	.33**	.22*	.00	-.15	.23*	.17	.74***	.74***	.
10. Age (months)	.46***	.34**	.44***	.09	.35**	.52***	.38***	.02	.27***
Mean	2.65	6.01	2.40	2.40	1.54	.00	2.56	3.19	.00
SD	1.59	2.46	1.37	1.64	.97	.58	.50	.55	.74
Cronbach's alpha	.48	.77	.42	.46	.	.	.35	.85	.
Split-half reliability	.53	.74	.33	.46	.	.	.35	.86	.

+ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$

Composite Scores. A rational thinking composite score was created by averaging available z scores from the denominator neglect, belief bias syllogism, base rate sensitivity, and otherside thinking tasks (as done by Toplak et al., 2014a; 2014b). We similarly created a normative thinking disposition composite from the NFC and AOT scales.

Results & Discussion

To what extent did the CRT-D function similarly to established measures of cognitive reflection? To answer this question, we first investigated adults' performance. Of particular interest was whether the CRT-D demonstrated similar relationships to the rational thinking and thinking disposition measures as the CRT and CRT-Alt. Next, we investigated whether children's performance on the CRT-D replicated well established relationships between cognitive reflection, rational thinking, and thinking dispositions. Finally, we examined the extent to which the CRT-D was a unique predictor of children's rational thinking and thinking dispositions when age was accounted for.

Relations Between the CRT-D, Rational Thinking, and Thinking Dispositions: Adults

Table 2 presents the means, standard deviations, reliabilities, and correlations among variables for adults. Hoeger's (2013) corrected version of Stieger's z-test of differences in dependent correlations revealed the correlation between the CRT-D and original CRT was similar to the correlation between the CRT-Alt and original CRT, $Z_H = -.68$, $p = .498$. Thus, the CRT-D was just as predictive of the original CRT as the CRT-Alt.

As seen in Table 2, all three measures of cognitive reflection were positively correlated with denominator neglect, base rate sensitivity, belief bias syllogisms, and the overall rational thinking composite. Only the CRT-Alt yielded a significant correlation with one of the thinking disposition measures (i.e., AOT). Corrected Stieger's z-tests revealed no differences between the CRT-D and original CRT in the strength of their correlations with the other variables. However, the CRT-Alt was more strongly correlated with responses on belief bias syllogisms and AOT than the CRT-D, $Z_H = -2.06$, $p = .039$, and $Z_H = -2.38$, $p = .017$, respectively.

Taken together, these results suggest the CRT-D functioned similarly to established measures of cognitive reflection, particularly the original CRT, at least with respect to the developmental versions of rational thinking tasks used in the present research.

Relations Between the CRT-D, Rational Thinking, and Thinking Dispositions: Children

Table 3 presents the means, standard deviations, reliabilities, and correlations among variables for children. Similar to adult measures of cognitive reflection (e.g., Toplak et al., 2014a; Thomson & Oppenheimer, 2016), the CRT-D was significantly correlated with children's denominator neglect, base rate sensitivity, AOT, and composite rational thinking and thinking disposition scores. The CRT-D also yielded positive correlations with children's otherside thinking and NFC (*r*'s = .19), though these were not significant using two-tailed tests. Given previously observed positive correlations between the CRT and these measures, one-tailed tests would also be appropriate.

Surprisingly, the CRT-D yielded a near-zero correlation with belief bias syllogisms, even though this relationship was present in our adult sample and has been observed in many prior studies with adults (e.g., Baron et al., 2015 Thomson & Oppenheimer, 2016; Toplak et al., 2014a). Also surprising was the lack of correlation between belief bias and any other measure, with the exception of the rational thinking composite (to which it contributed some of the variance). Given belief bias is typically a strong predictor of other rational thinking tasks and thinking dispositions for both children (Toplak et al., 2014b) and adults (e.g., Baron

Table 4. Regression analyses using age and CRT-D to predict rational thinking and normative disposition composites.

		Dependent variable					
		Rational thinking composite z-score			Normative disposition composite z-score		
Model	Predictor	R^2	β	95% CI	R^2	β	95% CI
1	Age (months)	.274	.012***	[.008, .167]	.073	.008*	[.002, .015]
2	Age (months)	.311	.010***	[.005, .015]	.125	.004	[-.003, .012]
	CRT-D		.078*	[.008, .148]		.118*	[.013, .223]

* $p < .05$; ** $p < .01$; *** $p < .001$

et al., 2015; Thomson & Oppenheimer, 2016), we speculate that something about our presentation format may have altered children's belief bias performance. For example, in comparison to individually completing items via paper-pencil assessment, our task involved a research assistant reading items aloud.

Overall, the CRT-D was broadly predictive of rational thinking tasks and thinking dispositions that are commonly associated with cognitive reflection in adults. These findings imply that the CRT-D is a valid measure of children's cognitive reflection. However, it is important to note that children's age yielded positive and significant correlations with the CRT-D (see Figure 1) and nearly every other variable. We thus ran further analyses, controlling for age.

Child CRT-D and Age Regression Analyses

To more closely examine the relationships between the CRT-D, children's age, and children's thinking, we performed hierarchical regressions for each composite score (Table 4). For each analysis, we first fit a model with age as the only predictor, and then a second model with both age and the CRT-D as predictors. All models satisfied the major assumptions of linear regression. After controlling for age, the CRT-D further predicted children's rational thinking composite score and explained an additional 3.7% of the variance, $F(1,90) = 4.84$, $p = .030$. Similarly, the CRT-D further predicted children's normative thinking disposition composite score and explained an additional 5.2% of the variance, $F(1,84) = 5.00$, $p = .028$. Thus, the CRT-D was predictive of both children's rational thinking and thinking dispositions, and its contributions were not fully explained by shared variance with age.

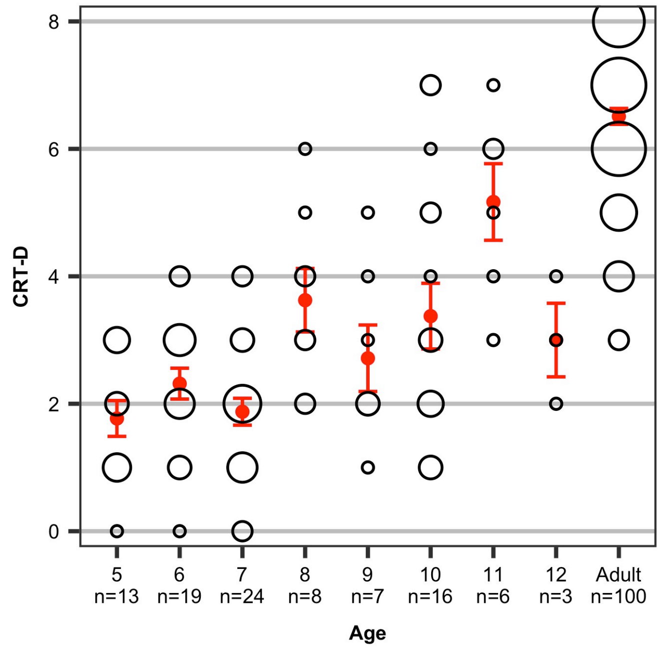

Figure 1. Mean (± SE) CRT-D scores by age.

General Discussion

In this study we developed and tested a new set of questions to measure cognitive reflection in children, the CRT-D. There is good reason to believe the CRT-D measures the same construct as the original CRT. For adults, the CRT-D predicted performance on the same measures as the original CRT (i.e., denominator neglect, base rate sensitivity, belief bias syllogisms, and alternative CRT items). Furthermore, the strengths of correlations between the CRT-D and these measures were highly similar to those yielded by the original CRT. For children, the CRT-D replicated previously observed correlations of the original CRT with denominator neglect, base rate sensitivity, otherside thinking, actively open-minded thinking, and need for cognition (e.g., Toplak et al, 2014a). Moreover, the CRT-D predicted children's overall rational thinking and thinking dispositions above and beyond children's age. Taken together, these results support the CRT-D as a valid measure of cognitive reflection in school-age children.

A critical question for future work is the extent to which children's cognitive reflection (as measured by the CRT-D) is predictive of thinking and reasoning in domains other than rational thinking (Pennycook et al., 2015). For example, Shtulman and McCallum (2014) examined the role of cognitive reflection in achieving conceptual change in six domains of science (astronomy, evolution, geology, mechanics, perception, and thermodynamics). They found the CRT explained more variance in college students' science understanding than STEM coursework, statistical reasoning ability, and nature of science understanding combined. Whether the CRT-D would predict children's conceptual change in science after controlling for other measures of cognitive ability, such as executive function, is an open question. If children's cognitive reflection is continuous with adults' cognitive reflection, a measure of the construct could be of great use in science education, as well as several other domains associated with cognitive reflection in adults (e.g., moral reasoning, evidential reasoning, religious cognition).

The CRT-D presented in the present research should be considered provisional. We took a conservative approach and did not eliminate items unless they failed to match the response structure of the original CRT (i.e., generating either a correct response or an intuitive lure response but not a variety of other, random responses). However, the relatively poor internal consistency of the CRT-D observed in the present study suggests some items may not be measuring the same construct, at least across the targeted age range. For example, Question 6 ("What weighs more, a pound of rocks or a pound of feathers?") demonstrated floor effects across our entire age range. While this is in some sense a feature of the original CRT, it may be the case that that some younger children simply lacked the knowledge required to answer correctly as opposed to failed to override the intuitive lure. More work with larger samples will need to be done to better understand the psychometric properties of the scale and individual item functioning across the targeted age range (e.g., Keller et al., 2016; Primi et al., 2016). That said, the present set of items was surprisingly effective at predicting children's rational thinking and thinking dispositions, even after children's age was accounted for.

To conclude, the CRT is a predominant measure of individual differences in analytic vs. intuitive thinking and meaningful predictor of a diverse range of psychological and behavioral outcomes. Here, we were largely successful in developing the CRT-D, a cognitive reflection test for children. This measure may prove useful in examining the role of analytic vs. intuitive thinking in a number of important domains of child cognition, such as scientific, moral, and social reasoning. Furthermore, the CRT-D will allow future researchers to investigate the developmental trajectory of cognitive reflection and its response to interventions during childhood.

Acknowledgements

We would like to thank the James S. McDonnell Foundation for supporting this research with an Understanding Human Cognition Scholar Award to Andrew Shtulman. We would also like to thank Thomas Egan, Isabel Geddes, Jasper Laca, and Precila Ortiz for their assistance with data collection and data analysis.

References

Baron, J., Scott, S., Fincher, K. S., & Metz, S. E. (2015). Why does the Cognitive Reflection Test (sometimes) predict utilitarian moral judgment (and other things)? *Journal of Applied Research in Memory and Cognition, 4*, 265–284.

Frederick, S. (2005). Cognitive reflection and decision making. *The Journal of Economic Perspectives*, *19*, 25–42.

Haran, U., Ritov, I., & Mellers, B. A. (2013). The role of actively open-minded thinking in information acquisition, accuracy, and calibration. *Judgment and Decision Making*, 8, 188–201.

Hoerger, M. (2013). Z_H: An updated version of Steiger's Z and web-based calculator for testing the statistical significance of the difference between dependent correlations. Retrieved from http://www.psychmike.com/dependent_correlations.php

Keller, U., Strobel, A., Wollschläger, R., Greiff, S., Martin, R., Vainikainen, M.-P., & Preckel, F. (2016). A Need for Cognition scale for children and adolescents. *European Journal of Psychological Assessment*, X. 1–13.

Kokis, J. V., Macpherson, R., Toplak, M. E., West, R. F., & Stanovich, K. E. (2002). Heuristic and analytic processing: Age trends and associations with cognitive ability and cognitive styles. *Journal of Experimental Child Psychology, 83,* 26–52.

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

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*, 453–469.

Shtulman, A., & McCallum, K. (2014). Cognitive reflection predicts science understanding. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.) *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 2937–2942). Austin, TX: Cognitive Science Society.

Stanovich, K., West, R., & Toplak, M. (2016). The rationality quotient: Toward a test of rational thinking. Cambridge: MIT Press.

Thomson, K. S., & Oppenheimer, D. M. (2016). Investigating an alternate form of the cognitive reflection test. *Judgment and Decision Making, 11*, 99–113.

Toplak, M. V., 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. (2014a). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking & Reasoning, 20*, 147–168.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014b). Rational thinking and cognitive sophistication: Development, cognitive abilities, and thinking dispositions. *Developmental Psychology, 50*, 1037–1048.

Travers, E., Rolison, J. J., & Feeney, A. (2016). The time course of conflict on the cognitive reflection test. *Cognition, 150*, 109-118.