Acta Psychologica 226 (2022) 103562

Contents lists available at ScienceDirect

Acta Psychologica

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

Does cognitive reflection predict attentional control in visual tasks?

Alessia Dorigoni [a,*], Jason Rajsic [b], Nicolao Bonini [a]

[a] Department of Economics and Management, University of Trento, Italy
[b] Department of Psychology, Northumbria University, United Kingdom of Great Britain and Northern Ireland

A R T I C L E I N F O

Keywords:
Cognitive reflection
Visual attention
Eye movements
Antisaccade
Visual search

A B S T R A C T

The cognitive reflection test (CRT) measures the ability to suppress an intuitive, but incorrect, answer that easily comes to mind. The relationship between the CRT and different cognitive biases has been widely studied. However, whether cognitive reflection is related to attentional control is less well studied. The aim of this paper is to investigate whether the inhibitory component of the CRT, measured by the number of non-intuitive answers of the CRT (Inhibitory Control Score), is related to the control of visual attention in visual tasks that involve overriding a bias in what to attend: an anti-saccade task and a visual search task. To test this possibility, we analyzed whether the CRT-Inhibitory Control Score (CRT-ICS) predicted attention allocation in each task. We compared the relationship between the CRT-ICS to two other potential predictors of attentional control: numeracy and visual working memory (VWM). Participants who scored lower on the CRT-ICS made more errors in the "look-away" trials in the anti-saccade task. Participants who scored higher on the CRT-ICS looked more often towards more informative color subsets in the visual search task. However, when controlling for numeracy and visual working memory, CRT-ICS scores were only related to the control of visual attention in the anti-saccade task.

1. Introduction

1.1. *Cognitive reflection*

Cognitive reflection is defined as the tendency to reflect on a question instead of reporting the first, potentially erroneous, response that comes to mind. It was initially measured by the Cognitive Reflection Test: the popular three item test in which people have to solve three problems that automatically generate intuitive responses (Frederick, 2005) and that require, in order to be solved, the inhibition of a heuristic response. All the questions in the CRT evoke a response that is incorrect but immediate and intuitive. This ability "to resist reporting the response that first comes to mind" (Frederick, 2005, p.35) has been interpreted in the literature using dual-system theories (Epstein, 1994; Evans, 2009; Kahneman & Frederick, 2002; Sloman, 1996; Stanovich & West, 2000). Errors on the CRT are believed to occur because System 2 (slow, effortful and reflective) fails to monitor System 1's outputs and to override its functioning (quick, intuitive and heuristic). As such, performance on the CRT is generally taken to indicate the degree to which a given person relies on System 1 in their thinking (Toplak et al., 2011).

Cognitive reflection has been shown to predict rational thinking and reasoning ability across a wide range of heuristics and biases (Toplak et al., 2011; Toplak et al., 2014), decision-making skills, time and risk preferences (Frederick, 2005), and thinking dispositions (Mata et al., 2013), which in turn are related to economic behavior and decision. For example, Noori (2016) showed that people with lower cognitive reflection are significantly more likely to exhibit the conjunction fallacy, illusion of control, overconfidence, base rate fallacy, and conservatism.

The relationship between cognitive reflection scores and a susceptibility to a wide range of heuristics and biases is consistent with a common cause of over-reliance on System 1. While the connection between cognitive reflection and heuristic thinking has been quite well-established, less is known about whether cognitive reflection is related to heuristic control of attention in perceptual tasks. Although perceptual processing is often described as automatic (Kahneman & Frederick, 2002) it can be highly contingent on cognitive control, as enabled by attention, which plays an important role in determining what information we gather from the environment (Simons & Rensink, 2005). Two-systems views (stimulus-driven and goal-driven) have a long history in research on visual attention (e.g. Jonides, 1981; Shiffrin & Schneider, 1977). For example, researchers have debated for decades to what extent top-down factors can override the influence of visual salience in

* Corresponding author.
E-mail address: alessia.dorigoni@unitn.it (A. Dorigoni).

https://doi.org/10.1016/j.actpsy.2022.103562
Received 24 July 2019; Received in revised form 24 February 2022; Accepted 14 March 2022
Available online 24 March 2022
0001-6918/

determining what is attended (Egeth et al., 2010; Folk et al., 1992; Luck et al., 2021; Theeuwes et al., 2006). To the extent that cognitive dispositions measured by the CRT reflect a general tendency towards heuristic control, they may not only affect how we think, but what we attend to when impulses and goals conflict. Whether or not cognitive reflection is predictive, of heuristic control of attention is important as many real-world behaviors depend on the ability to exert flexible control over visual processing (Kramer et al., 2019).

Attentional control is crucial in tasks that involve overriding a bias in what to attend, where the stimulus that should be attended is not the most salient stimulus, or not the stimulus that is intuitively relevant, such as the anti-saccade task (Hallett, 1978) where participants have to not look at a stimulus that appears on the screen. In the present study, we measured how our participants allocated their attention in two such tasks: an anti-saccade task and a visual search task.

1.2. Antisaccade task

In the anti-saccade task (Hallett, 1978) participants must overcome the urge to look at a salient stimulus that suddenly appears, and instead execute a saccade in the opposite direction (Munoz & Everling, 2004). Successful anti-saccade performance thus involves inhibiting the reflex to look at sudden onsets targets before looking in the opposite direction (Burnham & Anderson, 2004). This ability is poorer in the presence of some clinical disorders (Everling & Fischer, 1998) and cognitive load (Roberts et al., 1994). If differences in cognitive reflection reflect a broad reliance on heuristic processing, then lower CRT scores should be associated with more errors on anti-saccade trials, where visual heuristics conflict with goals. Such a relationship could reflect a common reliance on inhibition, as non-intuitive responses in the CRT require stopping the first intuition and correct responses on anti-saccade trials require the ability to inhibit an automatic response to look at the stimulus that appears.

1.3. Visual search task

In the visual search task, designed by Rajsic et al. (2015), participants are asked to decide if a target letter is, or is not, a specific color. On any given trial, the proportion of letters in each of two colors can vary. The goal of the task is to measure whether the arbitrarily chosen color used in trial instructions biases attention towards letters of that color. Because the target letter is always present, attending the letters in either the "yes" color or the "no" color exclusively would provide all the information necessary for an answer. The task affords two color-based search strategies (see Fig. 1). The first is to simply attend letters with the color mentioned in the question, regardless of the color proportion. The second is to search so as to minimize the number of letters searched (i.e., look at whatever color is the most rare, and infer the color by exclusion if the target is not in this set).

Even though the color asked about does not predict the color of the target overall, Rajsic et al. (2015) found that it produces an attentional bias to letters of this color. Although this inflexibility causes more searching than is strictly necessary, it presumably reflects the influence of a fast and simple search heuristic (matching bias; Rajsic, Taylor et al., 2017). Rajsic et al. (2015) interpreted this to mean that searchers were biased to seek confirmation of the question, providing a negative response when confirmation failed. Subsequent research has shown that this bias is caused by a tendency to rely on a template to guide attention (Rajsic, Wilson et al., 2017) instead of guiding attention to the rarest color, and is related to the need to provide an immediate yes/no response (Walenchok et al., 2020). That is, the matching bias seems to stem from reliance on a cognitively simple strategy, where the color that is attended is influenced by task settings that map the instructions to visual and motor elements.

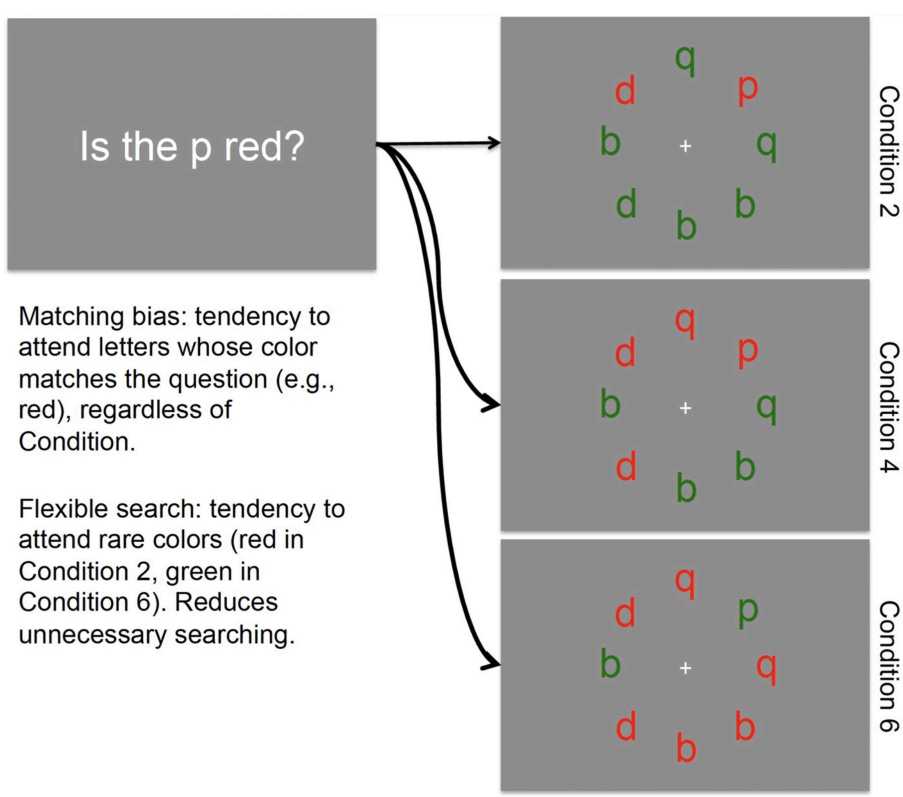

Fig. 1. An illustration of the search task with sample search prompt and displays (not to scale). Note that the target letter (in this case, "p") is always present, but can be either color (red/green) randomly. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

1.4. Visual working memory

In both the visual search task and the anti-saccade task, the degree to which participants can exert goal-driven control over attention is likely to be related to visual working memory capacity (Luck & Vogel, 2013). Visual working memory (VWM) capacity has been found to be quite variable between individuals, and this variance predicts aspects of attentional control (Fukuda & Vogel, 2009; Unsworth et al., 2014; Vogel et al., 2005). Furthermore, span measures of working memory capacity predict anti-saccade error rate (Unsworth et al., 2004) and selecting the more efficient sub-set in a visual search (Sobel et al., 2007). Therefore, as an additional variable of interest, we measured the visual working memory capacity of participants using a color change detection task.

At broader level, span measures of working memory have also been found to predict CRT performance (Stupple et al., 2013), as have additional measures of working memory capacity (Toplak et al., 2011). Such measures of working memory have also been shown to be important to the process of analytic thinking Bacon et al. (2008) and intelligence Kyllonen and Christal (1990). For these reasons, measuring VWM may provide a means of accounting for relationships between CRT performance and attentional control not due to cognitive reflection (as outlined in the next section). Although research suggests that visual change detection measures of VWM are related to intelligence like span measures of WM are (Shipstead et al., 2015; Unsworth et al., 2014), it is nonetheless important to consider that it is composite span scores that have been shown to be predictive of CRT performance (Stupple et al., 2013) rather than a single VWM capacity measure.

1.5. CRT and numeracy test

Empirically establishing whether cognitive reflection is related to the control of visual attention is complicated by recent concerns over whether the CRT solely measures cognitive reflection, or whether it measures other psychological constructs.

Although the empirical relationship between CRT scores and judgment and decision biases is strong, several authors have questioned the nature of the CRT, and its association with those biases (Weller et al., 2013; Welsh et al., 2013). On the one hand, cognitive reflection test scores may capture individual differences in general cognitive ability (i.e., intelligence or working memory; Blacksmith et al., 2019). It has also been argued that the CRT is mostly measuring participants' numeracy skill because of the high numerical component of its items (e.g., the famous bat-and-ball question about the correct price of a bat). There is currently an open debate on whether the CRT is merely a measure of a numeric ability and/or something different (at least in part). It has been suggested that the problem solving process during the CRT is a two-step process that implies blocking the intuitive response and getting the correct answer (Baron et al., 2015; Finucane & Gullion, 2010; Liberali et al., 2012; Sinayev & Peters, 2015; Szaszi et al., 2017). Whereas inhibitory control is crucial for the first step, numeric ability is essential for the second one; a disposition to engage in further reflection is necessary but not sufficient to lead to the correct response. The propensity "to devote cognitive resources to a task coupled with the right mindware may not be enough to find the correct answer if a participant has insufficient cognitive resource to reach the correct or normatively sanctioned conclusion" (Stupple et al., 2013) so the processes underlying CRT performance may include both cognitive reflection and numeric ability.

Campitelli and Gerrans, using a mathematical modeling approach based on a sample of 2019 participants, have showed that "a model that includes an inhibition parameter (i.e., the probability of inhibiting an intuitive response), as well as a mathematical parameter (i.e., the probability of using an adequate mathematical procedure), fitted the data better than a model that only included a mathematical parameter" (Campitelli & Gerrans, 2014), p. 434). This provides some evidence that CRT scores reflect a mixture of cognitive reflection and cognitive or numerical ability. For this reason, we will also consider how measures of numeracy (Weller et al., 2013) and VWM capacity (which we expect to be related to cognitive ability; (Shipstead et al., 2015) predict attentional control in a similar manner to CRT scores.

Several authors have proposed a different coding scheme of CRT responses to produce a better measure of the inhibition component (Böckenholt, 2012; Erceg & Bubić, 2017; Sinayev & Peters, 2015; Thomson & Oppenheimer, 2016). For example Sinayev & Peters, 2015 argue that coding the CRT responses into two categories (intuitive responses and non-intuitive responses) allows the separation of cognitive reflection from numeric ability; correct items suggest inhibition and numerical ability while giving an incorrect, intuitive answer is affected by whether or not someone uses cognitive reflection.

The former can be captured by counting the number of non-intuitive responses. Following these authors, we decided to use the Inhibitory Control Score in order to better measure cognitive reflection, distinct from numeracy. As suggested by those authors, we calculated the number of non-intuitive answers divided by the total number of answers (from here onwards, it is referred to as "CRT-ICS"). However, for completeness, statistical analyses run on the traditional CRT scoring method (right answers: the number of correct answers divided by the total number of answers) are also reported in Appendix A.

Since CRT was first published (Frederick, 2005), an increasing number of scholars have started to use a so called "extended CRT" where other cognitive reflection tests are added to the original three-item CRT (Toplak et al., 2014; Stagnaro et al., 2019; Weiss et al., 2021). There are many reasons for this practice. First of all, the familiarity with the Frederick's test. Due to the popularity of the CRT, many participants may already know the right answers (Haigh, 2016). This compelled researchers to devise new items measuring the same construct that were eventually added to the original test. For example, Stieger and Reips (2016) suggest "to use more recent multi-item CRTs with newer items" in order to cope with the ceiling effect. Second, the extended CRT has a more normal distribution and lower floor effect than the original CRT (Stagnaro et al., 2019). Another reason to use an extended CRT is related to the numeracy confound previously-mentioned. For example, several authors (e.g. Pennycook & Rand, 2019; Stagnaro et al., 2019) added the Thomson and Oppenheimer (2016) cognitive reflection test which is based on verbal items which are less correlated with numeracy. Following this common practice, we decided to collect data from four different versions of the CRT, in a randomized order (Frederick, 2005; Primi et al., 2016; Thomson & Oppenheimer, 2016; Toplak et al., 2014); we removed the original Frederick's questions from the variant versions to prevent repetition.

1.6. Hypothesis

In the present paper, we tested the hypothesis that cognitive reflection, as measured by CRT-ICS scores, can predict how participants allocate their visual attention in the anti-saccade task (Hallett, 1978) and in a visual search task Rajsic et al. (2015). For each task, we hypothesized that:

Anti-saccade task

HP_1: People with lower CRT-ICS will make more errors in the anti-saccade trials than people with higher CRT-ICS. This is due to their inability to stop an automatic response that drives them to look at a target that appears.

On the other hand, we expected that the CRT-ICS would not predict pro-saccade performance because in that specific task there is not an automatic response to block. As for numeracy we expected that since there are not numerical components in this kind of task, numeracy would not predict the number of errors on anti-saccade trials, unlike CRT-ICS scores. For VWM, we predicted that if the CRT-ICS scores only predict attentional control due to shared variance with working memory

(Stupple et al., 2013; Toplak et al., 2011), then VWM would predict the rate of anti-saccade errors as well as, or better than, CRT-ICS scores.

Visual search task

$HP_{2.1}$: People with lower CRT-ICS will exhibit the matching bias over all color set sizes, with their first fixations being most likely to go to a letter whose color matches the rule (the one mentioned in the rule).

$HP_{2.2}$: People with higher CRT-ICS will guide attention to the color that provides the most information (i.e., lets them finish a search with fewer letters being identified). This means that when there are six letters with the matching color and two letters with the mismatching color, these participants will be less biased (i.e., more likely to attend "green" in Condition 6, Fig. 1).

As for numeracy we expected that since there are not numerical components in this kind of task, numeracy would not be able to predict a change in matching bias in Condition 6. For VWM, we predicted that it could be related to how well participants can fixate a task relevant color overall, but that it would likewise not predict a change in matching bias in Condition 6.

2. Methods

2.1. Participants

54 university students from the local University (34 female, 20 male; mean age = 23.6 years, SD = 4.5) participated to this study. We collected as many participants as our resources allowed to increase statistical power; during the instructions we told participants that they will receive 3 euros for participation and between 3 and 11,5 euros on the basis of their performance in one of the tasks.

2.2. Materials

The experiment was run on a Dell computer with a 23 inch screen. Responses were collected with a standard USB keyboard. Eyetracking data was collected using an Eye Link 1000 Plus (SR Research) Binocular Tower Mount which provides a data acquisition at up to 2000 Hz, sampling from the right eye. Computer-based tasks were programmed in Matlab using the Psychophysics Toolbox (Brainard & Vision, 1997; Kleiner et al., 2007; Pelli, 1997). CRTs and the numeracy test were collected using paper-and-pencil.

2.3. Stimuli and procedure

Each participant, after reading the instructions, completed the tasks on the computer with eye tracking. In addition to the anti-saccade task (Hallett, 1978), visual search task (Rajsic et al., 2015) and visual working memory task (Luck & Vogel, 1997), participants completed three other tasks as part of a larger research project. After this experiment section, they completed the four different CRTs mentioned above and the numeracy test in a random order, without any time limit, using paper-and-pencil.

2.4. Anti-saccade task

This task consisted of 80 trials; 40 pro-saccade trials and 40 anti-saccade trials, randomly intermixed. Each trial started with the instruction "Look at" or "Look away" printed in the center of the screen for 2 s to indicate what the appropriate response would be Fig. 2; a participant is asked to make a saccade in the direction away from the stimulus (anti-saccade condition) or a saccade in the direction of the stimulus (pro-saccade condition). This screen was replaced by a fixation cross and two small, white rectangles (0.3°), 9° to the left and right of fixation. After a randomized delay (evenly sampled from 2 s to 2.5 s), a 1.2° target rectangle at one of the two locations (randomly selected)

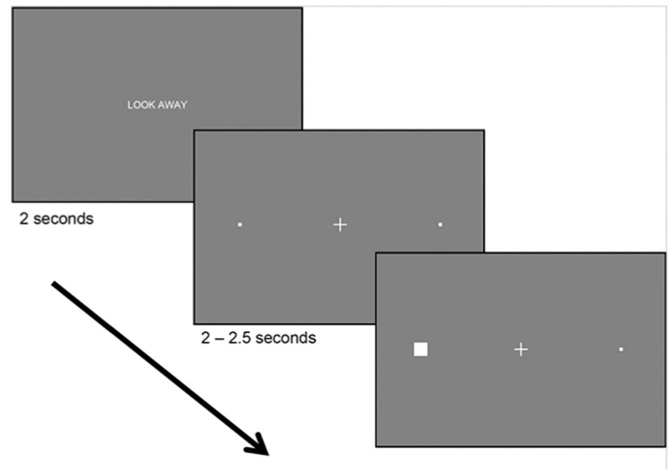

Fig. 2. An illustration of an anti-saccade trial. Note that the target flickered on onset.

flickered 3 times (50 ms.

on, 50 ms off) as shown in Fig. 2.

When participants' gaze moved 3° left or right of fixation, this was considered a response. After 300 ms, the next trial began unless an error was made. Errors were followed by 1 s of feedback ("You looked the wrong way!"). On trials where participants moved their eyes before the target, trials ended abruptly, and error feedback was presented for 2 s ("Do not move your eyes before the target occurs."). All such false-start trials were recycled so that every participant provided 80 trials with valid target responses.

2.5. Visual search task

This task consisted of 90 trials containing a random mixture of 15 trials in each of six conditions, fully crossing two factors (matching set size [2, 4, or 6 letters] and correct response [yes, no]). Each trial began with a 1-second blank screen, after which the search prompt was given. Search prompts contained that trial's target letter and a color (e.g., "Is the p red?"), presented for 2 s. Target letters were randomly selected from the set [p, q, d, b], with the remaining three letters used as distractors, and colors were randomly selected from the set [red, green], such that either 2, 4, or 6 letters matched the color in the prompt and the target letter either was (for "yes" trials) or was not (for "no" trials) the color from the prompt. After the prompt, participants saw a 500 ms screen with just a fixation mark in the center of the screen. After this, the search display appeared, with eight letters (seven distractors and one target letter, approximately 0.6° X 1.1°) appearing evenly distributed on the circumference of a 10.5° radius imaginary circle centered on the fixation mark (Fig. 1).

This display remained on screen until the participant responded by pressing either the "z" key for "yes" or the "m" key for "no". Immediately following responses, the word "correct" or "incorrect" appeared in the center of the screen for 1 s.

We analyzed the first fixation because it provides a measure of initial attentional allocation without needing to consider how participants integrate information during search (Rajsic et al., 2015).

2.6. Visual working memory

To measure VWM capacity, we used a color change detection task (Luck & Vogel, 1997). This task consisted of 160 trials, fully crossing two factors (number of colors [2, or 6], and change [absent, present]. Each trial began with a one-second fixation mark on a blank screen. The memory sample display then appeared for 100 ms, consisting of either 2 or 6 colored squares, randomly placed in an imaginary square grid spanning 9.7°, centered on fixation (Fig. 3).

Squares were 1.2° in height and width, separated by at least 2.4° (center to center). This display was replaced by a blank screen with a fixation mark for 900 ms. After this, a memory test display appeared until a response was entered. On change absent trials, the test display was identical to the sample display. On change trials, one square's color was replaced with an unused color. Colors were selected from a set of nine highly discriminable colors (red, green, blue, magenta, yellow, orange, gray, white, and black). Participants reported changes with the "A" key, and no change with the "L" key. A break was provided every 60 trials.

Memory capacity was calculated as k using the whole display equation (Rouder et al., 2011) with hit rate and false alarm rate at set size six only. This equation corrects for the possibility of guessing using false alarm data, and is computed as:

$$k = N\left(\frac{h-f}{1-f}\right)$$

where N is six (the number of colors to remember), and h and f are the measured hit-rates and false-alarm rates, respectively, for individual participants. Data from set size two were used simply to verify that participants used the correct response mapping, and to reverse responses when false alarm rate exceeded hit rate.

2.7. Cognitive reflection task

We collected data from four different versions of the CRT (Frederick, 2005; Primi et al., 2016; Thomson & Oppenheimer, 2016; Toplak et al., 2011) where the original three-items CRT were used only once.

The CRT-ICS score was measured as the number of non-intuitive answers divided by the total number of questions. In Appendix A we report the results for the CRT classic score coding (the number of correct answers divided by the total number of answers).

All the answers were collected using paper-and-pencil and they were presented in a random order.

2.8. Numeracy

As a test of mathematical and probabilistic reasoning we used the numeracy scale developed by Weller et al., 2013. This scale consisted of eight items, five from the original Lipkus et al. scale, one of the Peters et al. items and two from the CRT scale by Frederick. Because Frederick's questions are already included in the CRT test, we decided to remove them from the numeracy set and we used the 6-items numeracy test.

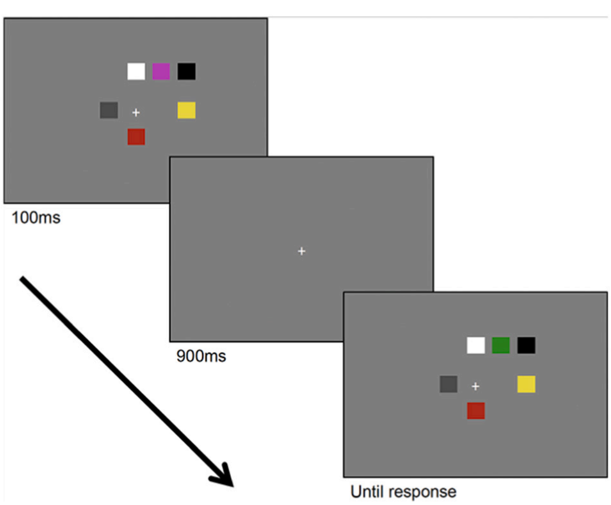

Fig. 3. An illustration of the memory task, showing a "change" trial.

3. Results

3.1. Task analysis

A preliminary analysis between the two kinds of CRT codings (proportion of non-intuitive answers, CRT-ICS, and the number of right answers, Classic CRT, over the total number of answers; Fig. 4) shows that in Thomson and Oppenheimer test most of the time when a non-intuitive answers was given this is also correct. There is therefore an overlap between the CRT-ICS and the Classic CRT score coding.

Since the aim of this paper is to understand the role of the inhibitory component of the CRT-ICS on visual attention, we first analyzed if we should use the Inhibitory Control Score of the 14-item CRT composite (14-item CRT-ICS) or if we should separate the four questions of the test by Thomson and Oppenheimer (2016; CRT-ICS 2) because in the literature it appears to be a purer variable and to rely less on numeracy than the original CRT. For this reason we analyze the internal consistency of these measures and the correlations among them.

In order to understand the internal consistency of these measures we used a special case of Cronbach's α computed for dichotomous scores (intuitive/non-intuitive answer), the Kuder-Richardson formula 20 (KR-20).

As we can see from the Table 1 there are high and significant correlations among the two CRT tests, and the composite 14-item CRT-ICS.

17% of the participants received a score of four in the 4-item of the CRT-ICS 2 and a mean of 0.6 (SD = 0.27, KR-20 = 0.61). The 14-item CRT-ICS scale had a more normal distribution, with only with only 0.1% scoring a full fourteen, a mean of 0.67 (SD = 0.19, KR-20 = 0.67). For the considerations mentioned above, we decided to include the CRT-ICS 2 along with the other CRT items to obtain the 14-item CRT-ICS composite.

As we can see from Table 2, the mean of the 6-item numeracy is 0.68 (SD = 0.25) with a significant Spearman correlation with the CRT-ICS = 0.64 (the correlation with the classic CRT score coding = 0.70); numeracy has a lower correlation with the CRT-ICS compared to the Classic CRT.

The average k in the visual working memory task (VWM) is 2.75 with a SD = 1.15 and a correlation with the 14-item CRT-ICS = 0.33.

3.2. Dependent variables: control of visual attention

As we can see from the Fig. 5 in the anti-saccade task (panel A), we found that participants made more errors on anti-saccade (M = 9.95%, SD = 7.2%) than pro-saccade (M = 4.2%, SD = 3.1%) trials, $t(53) = 6.09$, $p < .001$, confirming that anti-saccade trials led to impulsive responses. We also noticed that participants made a considerable number of saccades before the target appeared (which prematurely ended the trial). These false start trials were no more frequent in anti-saccade (M =

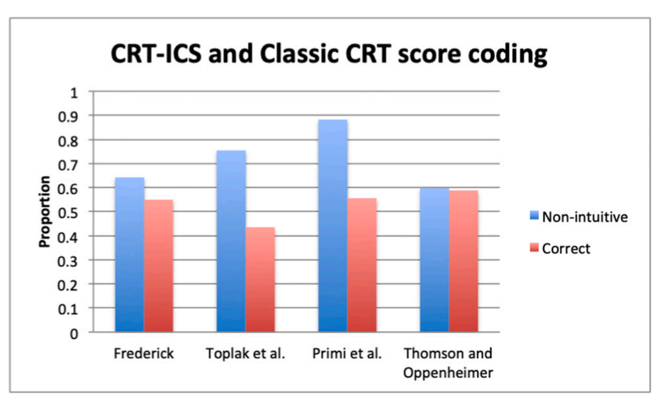

Fig. 4. Proportion of non-intuitive answers (CRT-ICS) and the number of right answers (Classic CRT score coding) over the total number of answers.

Table 1
Intercorrelation (Spearman) among CRT-ICS 2 and 10-item CRT-ICS.

	CRT-ICS 2	10-item CRT-ICS	14-item CRT-ICS
10-item CRT-ICS	0.55***		
14-item CRT-ICS	0.81***	0.94***	
Mean	0.6	0.7	0.67
SD	0.27	0.19	0.19
Reliability (KR-20)	0.61	0.51	0.67

* p < .1; ** p < .05; *** p < .01.

Table 2
Intercorrelation among Visual Working Memory, Numeracy, 14-item CRT-ICS (CRT-ICS) and 14-item Classic CRT (CRT).

	VWM	6 item numeracy	CRT-ICS	CRT
6 item numeracy	0.26			
CRT-ICS	0.33*	0.64***		
CRT	0.33*	0.70***	0.9***	
Mean	2.76**	0.68	0.53	0.71
SD	1.16	0.25	0.24	0.21

* p < .1.
** p < .05.
*** p < .01.

12.9%, SD = 10.9%) than pro-saccade (M = 12.9%, SD = 9.1%) trials, t(53) = 0.07, p = .95.

In the visual search task (panel B), as in previous studies the average correct search time was fastest for "yes" responses (red line), $F(1, 53) = 126.70$, $p < .001$, $\eta^2_p = 0.71$. Search was also slower with more question-matching colored letters, $F(2, 106) = 68.01$, $p < .001$, $\eta^2_p = 0.56$, with significant linear, $F(1, 53) = 95.60$, $p < .001$, $\eta^2_p = 0.64$, and quadratic, $F(1, 53) = 29.34$, $p < .001$, $\eta^2_p = 0.36$, trends. These factors also interacted, $F(2, 106) = 4.66$, $p = .011$, $\eta^2_p = 0.08$. The quadratic trend shows that participants did not exclusively attend to the "yes" colored letters, but were nonetheless biased to attend them given that searches were faster when targets appeared in the small, matching subset than when they appeared in the small, mismatching subset, t(53) = 12.57, p < .001. At all set sizes, search time for "yes" responses was faster than search time for "no" responses, ts > 4.86, ps < .001. These results generally replicate previous findings that participants' attention is biased by the particular color they are asked about, although this is indeed a bias and not rigid strategy.

3.3. Anti-saccade task and cognitive reflection

We used generalized linear mixed models (Bates et al., 2014), which incorporated both fixed-effects parameters and random effects (subjects), to evaluate the hypotheses mentioned. They allowed us to take advantage of trial-level data across all participants without collapsing all the data into sample averages.

To measure anti-saccade errors, we coded each trial's response for each participant as a 0 for correct and a 1 for incorrect saccades, excluding trials where eyes moved before the target onset.

We computed a generalized linear mixed model on both, pro-saccade trials ("look towards" trials) and anti-saccade trials ("look away trials"), to test our hypothesis.

As we can see from Table 3, the CRT-ICS did not predict the performance in the pro-saccade trials in model 1, without numeracy and visual working memory, nor in the model 4, with numeracy and visual working memory (SD in brackets). Neither numeracy (6-item Numeracy) nor visual working memory (VWM) has a predictive power in the models where they are single predictors either (model 2 and model 3).

In the anti-saccade trials, as we can see from the Table 4, CRT-ICS predicts how often participants made errors only in model (1), that is without numeracy and working memory.

The first model presents a lower AIC than all other models, and the negative beta shows that participants with lower CRT-ICS were more likely to make anti-saccade errors confirming hypothesis 1.

$$(\beta_{M1,CRT-ICS} = -0.998, z = -1{,}919, p = .054)$$

Adding numeracy and visual working memory to the model makes CRT-ICS not reach significance for a two-tailed test, but it still meets significance for a one tailed

$$(\beta_{M4,CRT-ICS} = -1.131, z = -1{,}662, p = .096).$$

However, it is important to note that these two variables, in separate models (model 2 for numeracy and model 3 for visual working memory) do not predict anti-saccade errors. We calculated the variance inflation

Table 3
Errors in the "look towards" trials in the anti-saccade task.

	Dependent variable			
	Errors in the "look towards" trials			
	(1)	(2)	(3)	(4)
CRT-ICS	−0.227*,**			−0.210
	(0.506)			(0.677)
6-Item Numeracy		−0.141		−0.033
		(0.419)		(0.551)
VWM			−0.005	0.006
			(0.090)	(0.093)
Constant	−2.963***	−3.029***	−3.110***	−2.971***
	(0.372)	(0.301)	(0.288)	(0.419)
Observations	2160	2160	2160	2160
Log Likelihood	−377.153	−377.197	−377.252	−377.149
Akaike Inf. Crit.	760.306	760.394	760.503	764.298

* p < .1.
** p < .05.
*** p < .01.

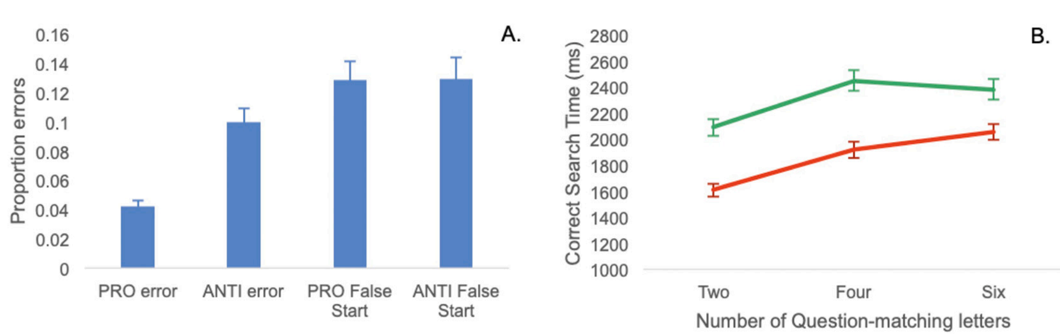

Fig. 5. Average performance in the anti-saccade task (panel A) and in the visual search task (panel B). In the panel B the red line represents the search time for "yes" responses and the green line represents the search time for "no" responses. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 4
Errors in the "look away" trials in the anti-saccade task.

	Dependent variable			
	Errors in the "look away" trials			
	(1)	(2)	(3)	(4)
CRT-ICS	−0.998*			−1.131*
	(0.520)			(0.681)
6-Item Numeracy		−0.381**		0.262
		(0.438)		(0.555)
VWM			−0.085	−0.049
			(0.091)	(0.092)
Constant	−1.637***	−2.089***	−2.092***	−1.574***
	(0.378)	(0.314)	(0.292)	(0.422)
Observations	2160	2160	2160	2160
Log Likelihood	−685.839	−687.256	−687.198	−685.599
Akaike Inf. Crit.	1377.677	1380.512	1380.396	1381.198

* $p < .1$.
** $p < .05$.
*** $p < .01$.

Table 5
False Start in the anti-saccade task.

	Dependent variable			
	False start			
	(1)	(2)	(3)	(4)
CRT-ICS	−1.117**			−0.479
	(0.543)			(0.711)
6-Item Numeracy		−1.110**		−0.891
		(0.443)		(0.576)
VWM			−0.008*	0.046
			(0.098)	(0.095)
Constant	−1.285***	−1.330***	−2.057***	−1.274***
	(0.400)	(0.316)	(0.316)	(0.440)
Observations	5049	5049	5049	5049
Log Likelihood	−1959.863	−1958.865	−1961.904	−1958.576
Akaike Inf. Crit.	3925.726	3923.729	3929.808	3927.152

* $p < .1$.
** $p < .05$.
*** $p < .01$.

factor (VIF) in order to test the multicollinearity and to remove any variable with a high VIF value from the model (values above 5 are commonly considered an index of excessive correlation among explanatory variables). These values showed that for the three predictors (CRT-ICS, 6-item Numeracy, and VWM) there were no indices higher than 5 (CRT-ICS = 2.4; 6-item Numeracy = 2.4 and VWM = 1.1).

If we compare the pro and anti-saccade trials using a median-split on CRT-ICS scores (Fig. 6) we can see that, in the pro-saccade trials, there is no difference between low and high CRT-ICS participants. Instead, there is a difference between these two groups in the anti-saccade trials, where it is more likely that people with lower CRT-ICS make errors.

As noted in Section 3.1, we found a surprising occurrence of false-starts. We considered these saccadic false starts as a form of impulsivity, and so we decided to analyze if CRT-ICS predicts false starts in the anti-saccade task. To the extent that participants make impulsive eye movements before the target onsets, these too should be more frequent in participants with lower CRT-ICS scores. In order to test this hypothesis we used as dependent variable whether the participant made a false start (1) or not (0) on each trial.

As can be seen in Table 5 CRT-ICS (model 1) and numeracy (model 2) predicted false starts in the model with these two variables as single variables. In the first and second model:

$$(\beta_{M1,CRT-ICS} = -1.12, z = -2.05, p = .03)$$

$$(\beta_{M2,NUMERACY} = -1.11, z = -2.50, p = .01)$$

The negative beta value shows that participants with lower CRT-ICS

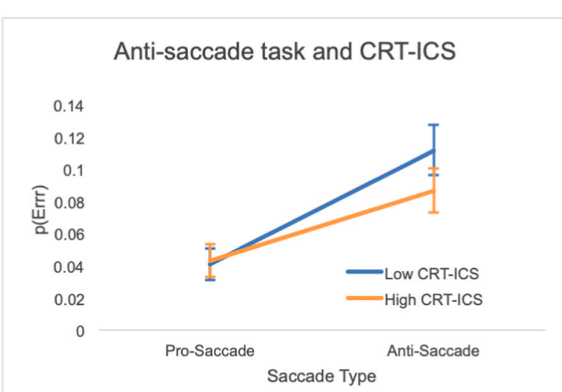

Fig. 6. Probability of making an error in pro-saccade and anti-saccade trials depending on median-split of CRT-ICS (95% confidence interval bars).Participants with the median value of CRT-ICS were considered "low" CRT-ICS.

or lower numeracy were more likely to make false starts. However, visual working memory capacity did not predict false starts in the anti-saccade task.

In the fourth model, the three variables have no predictive power. We found a VIF index of 2.52 for CRT-ICS, 2.50 for the numeracy test and 1.07 for visual working memory, which allows us to exclude collinearity problems between the variables. From these results, false starts appeared to be related to variance shared by CRT-ICS and numeracy score.

3.4. Visual search task and cognitive reflection

In order to test whether CRT-ICS scores predicted visual matching bias, we used as a dependent variable a binary measure of whether the first fixation of each correct trial was on a matching color (1) or if it was on a mismatching color (0). To first test whether cognitive reflection could be related to attentional control in this task, we considered a model (model 1) with the CRT-ICS as independent variable. Matching set size was entered as a dummy variable, with set size four as the intercept. This meant that the model was set up to predict increases or decreases in the probability of fixating on a matching color when the number of matching colors decreased to two (Condition 2), and when it increased to six (Condition 6). Then, to test whether numeracy or VWM could alternatively account for a relationship between the CRT-ICS scores and attentional control, we ran additional models (2) and (3) where numeracy and VWM, respectively, replaced CRT-ICS scores. Finally, to test which predictor (CRT-ICS, numeracy, or VWM) best accounted for differences in matching bias, we ran model (4), where each predictor variable was included as an interaction term with the matching set size dummy variable (Condition 2 and Condition 6). The results of individual differences in on matching bias are shown in Table 6.

As expected, the interaction between CRT-ICS and matching set size six ("Condition 6" in Model 1, Table 6) is significant

$$(\beta M\,2, Condition6{*}CRT-ICS = -0.811, z = -2.18, p = .03).$$

The beta coefficient of this interaction was negative, meaning that people with higher CRT-ICS reduced their attention to matching colors more often when matching colored letters provided less information, compared to people with lower CRT-ICS. This is consistent with our hypothesis that CRT-ICS scores predict better attentional flexibility in this task: those who scored higher in the CRT-ICS attended the color that reduced search load more often, from their first eye movement. The interaction between matching set size two ("Condition 2" in the table) and the CRT-ICS was not significant.

To test whether numeracy and VWM could predict matching bias like

Table 6
Matching bias in the visual search task.

	Dependent variable			
	Matching bias			
	(1)	(2)	(3)	(4)
Condition 6	1.278***	1.310***	0.989***	1.501***
	(0.276)	(0.225)	(0.208)	(0.309)
Condition 2	−0.730***	−0.433**	−0.679***	−0.773***
	(0.254)	(0.206)	(0.194)	(0.286)
CRT-ICS	0.370			−0.202
	(0.283)			(0.365)
Condition 6: CRT-ICS	−0.811**			−0.127
	(0.371)			(0.488)
Condition 2: CRT-ICS	0.335			0.605
	(0.345)			(0.455)
6-Item Numeracy		0.573**		0.624**
		(0.235)		(0.300)
Condition 6: 6-item Numeracy		−0.904***		−0.782*
		(0.312)		(0.403)
Condition 2: 6-item Numeracy		−0.090		−0.455
		(0.288)		(0.373)
VWM			0.080	0.065
			(0.049)	(0.050)
Condition 6: VWM			−0.097	−0.062
			(0.065)	(0.067)
Condition 2: VWM			0.063	0.053
			(0.060)	(0.062)
Constant	0.126	0.003	0.150	−0.083
	(0.209)	(0.168)	(0.156)	(0.229)
Observations	4742	4742	4742	4742
Log Likelihood	−3042.283	−3041.708	−3042.764	−3036.435
Akaike Inf. Crit.	6098.566	6097.416	6099.529	6098.870

* $p < .1$.
** $p < .05$.
*** $p < .01$.

the CRT-ICS did, we ran two additional models (Model 2 and Model 3), where numeracy scores and VWM scores, respectively, replaced CRT-ICS in the model. In these models, numeracy predicted a reduction in matching bias in Condition 6

$$(\beta_{M2,Condition6*NUMERACY} = -0.904, z = -2.89, p = .003)$$

similarly to CRT-ICS scores. VWM, on the other hand, was not related to matching bias in its model. These results suggest that it may be numeracy, and its associated ability to improve CRT scores, that is related to attentional flexibility in this search task. When all predictors were entered into the same model, only numeracy scores predicted a reduction in matching bias in Condition 6

$$(\beta_{M4,Condition6*NUMERACY} = 0.781, z = 1.94, p = .052).$$

So although our hypothesis was supported, it is important to consider that differences in numeracy, rather than cognitive reflection, predict attentional flexibility.

To illustrate the relationship between CRT-ICS scores and the tendency to look at matching stimuli, we computed a matching bias index (Rajsic, Wilson et al., 2017), which measures how much a participant was biased to look at the matching color (controlling for the chance). It is described by the following equation:

$$bias(p(match), chance) = \begin{cases} \frac{p(match) - chance}{1 - chance}, p(match) \geq chance \\ \frac{p(match) - chance}{chance}, p(match). < chance \end{cases}$$

When the measured probability of inspecting the template-matching color is greater than or equal to chance (0.25, 0.5, and 0.75 for the matching subset sizes 2, 4, and 6, respectively), the extent of the matching bias is given by p(match) minus chance divided by 1 - chance, which expresses the degree to which fixations go to matching-color letters above what would be expected by random eye movements. When the measured probability of inspecting the template-matching color is lower than the chance, the matching bias is given by p(match) minus chance divided by chance.

As we can see from the Fig. 7, all participants' matching bias reduced as the matching set-size increased. This is consistent with previous research (Rajsic, Wilson et al., 2017) and shows that matching bias is not the sole influence on attention in this task. However, dividing participants using a median split of CRT-ICS scores shows that participants with higher CRT-ICS scores show a larger difference in matching bias between set size 4 and set size 6, when it becomes more efficient to fixate a mismatching letter to decide what color the target is. In other words, high CRT-ICS participants seem to better able to use an attentional strategy that allows them to minimize the number of stimuli processed. Much the same is evident when participants are split by numeracy.

4. Discussion

Cognitive reflection is believed to result from individual differences in the control of System 2 processes over System 1 processes. The CRT was designed to exploit heuristics that lead to an erroneous but intuitive response, and scores on this task predict a wide range of heuristics and biases indicative of System 1 processes (Jackson et al., 2016). In the present study, we sought to test whether four combined cognitive reflection tests predict the allocation of attention in two visual tasks. To do this, we measured CRT scores for participants who completed an anti-saccade task, measuring the ability to resist attending to a salient input (Hallett, 1978; Munoz & Everling, 2004) and a visual search task, measuring the ability to flexibly change attentional settings on each trial (Rajsic et al., 2015). Both tasks pit heuristic actions (look at the target; attend the color mentioned in the question) against more controlled responses (look away from the target; look at the color that provides the most information). We found that CRT-ICS scores did predict flexible attentional control in these two tasks. This was true both when considering the likelihood of generating a non-intuitive response (CRT-ICS) and also when considering the likelihood of generating the correct response (CRT-classic coding; see Appendix A). However, we also found that numeracy predicted control over matching bias in the search task better than CRT-ICS scores did, and that including numeracy as a predictor of anti-saccade errors rendered CRT-ICS scores' predictiveness non-significant. This suggests that the shared variability between CRT-ICS and numeracy is likely what is important in our measures of attentional control.

In the anti-saccade task, we found that CRT-ICS scores were specifically predictive of more anti-saccade errors (Unsworth et al., 2004). The lack of a difference for pro-saccades suggests that it was specifically

Fig. 7. Matching bias index and CRT-ICS, using a median split (95% confidence interval bars).Participants with the median value of CRT-ICS were considered "low" CRT-ICS.

when participants needed to shift attention away from a salient onset that differences in CRT-ICS scores were related to task accuracy. Unexpectedly, we also found that participants made impulsive saccades (false starts) before target onset in both conditions. Although CRT-ICS scores predicted these too, in this case numeracy scores were also predictive of errors. This is unintuitive, as the need to wait for a task signal before executing a saccade does not involve numerical processing. It seems more likely that, in this case, differences in broader cognitive capacity that contribute to both CRT-ICS scores and numeracy account for this relationship. In light of previous connections between CRT scores and impulsivity (Graffeo et al., 2015), it may be better to consider false-starts as impatience, or temporal impulsivity. Frederick (2005) reported that higher CRT scores were related to differences in temporal discounting, and it may be that this relationship is due to underlying connections between numeracy and temporal preferences (Welsh et al., 2013; Dohmen et al., 2010).

In the visual search task, we tested whether CRT-ICS scores were predictive of control over visual matching bias (Rajsic et al., 2015; Rajsic, Taylor et al., 2017). When attending the matching color was useful (condition 2) or harmless (condition 4), both high and low-CRT-ICS participants showed matching bias (Fig. 7). However, when attending the matching color is less helpful (condition 6), higher CRT-ICS scores were associated with a larger drop in the probability of fixating on a matching colored letter. Although this was consistent with our hypothesis that greater cognitive reflection makes it easier to overcome matching bias when it is less helpful, we also found that numeracy scores predicted the same reduction in matching bias as CRT-ICS scores. Given that this task does not involve overt numerical components, it is not immediately clear why this would be the case. However, one reason may be because it is the different color proportions in the search display that provide the cue to whether the color mentioned in the question, or the color not mentioned, is the most informative about the correct response. Appreciating the different strategies these stimulus differences imply may rely on numerical intuitions, which also support probabilistic or proportional reasoning. For example, although visual aids (i.e., graphs) improve understanding of risk overall, those with higher numeracy show better understanding of risk when using visual aids than those with lower numeracy (Petrova et al., 2017). On the other hand, VWM did not predict differences in matching bias despite previous findings supporting this prediction (Fukuda & Vogel, 2009; Sobel et al., 2007; Unsworth et al., 2014).

From these results, we conclude that a relationship between cognitive reflection and attention is difficult to evidence. Whereas CRT-ICS scores were uniquely able to predict avoidance of incorrect saccades in the anti-saccade task, numeracy attenuated the relationship between CRT-ICS scores when included in the model. Moreover, the relationship between CRT-ICS scores and matching bias in visual search was equally well explained by numeracy. Finally, we observed a tendency to make an eye movement before pro and anti-saccade trials that was related to both cognitive reflection scores and numeracy. While it is possible to account for how numerical competency might relate to the matching bias, by helping to making the utility of visual proportions as a cue for attention more salient, it is not as easy to explain why numeracy would predict anti-saccade error rates. As such, it could be that more general differences in cognitive abilities between our participants are related to both their CRT scores and their numeracy and that these relationships reflect a broader willingness or ability to adopt more cognitively demanding, but performance-enhancing, states of control.

In light of the complexities of directly measuring underlying differences in cognitive re-flection, further research is needed in order to understand if there is a link between cognitive reflection and attentional control (and indeed if cognitive reflection is a unique construct; Attali & Bar Hillel, 2020; Erceg, Galić, and Ružojčić, 2020). In particular, it is possible that other variables (such as intelligence, working memory or other constructs we did not measure) may explain apparent relationships between cognitive reflection and attention. When considering attentional control, there are many visual tasks that involve some suppression of attention to salient, habitual, or intuitive stimuli (e.g., Anderson et al., 2011; Eriksen & Eriksen, 1974; Navon, 1977; Stroop, 1935; Theeuwes, 1992), and much work remains to understand to what degree such structurally common tasks rely on shared or distinct individual differences in cognitive abilities and dispositions (e.g. Draheim et al., 2020; Rey-Mermet et al., 2018).

Author's note

We would like to include a thank you to both of our (anonymous) reviewers for their suggestions and for their helpful discussions of the nature of cognitive reflection as a construct.

Appendix A. CRT classic coding

Table 7
Errors in the "look towards" trials in the anti-saccade task.

	Dependent variable			
	Errors in the "look towards" trials			
	(1)	(2)	(3)	(4)
Classic CRT	−0.403*,**			−0.616
	(0.437)			(0.629)
6-Item Numeracy		−0.141		0.261
		(0.419)		(0.594)
VWM			−0.005	0.017
			(0.090)	(0.092)
Constant	−2.915***	−3.029***	−3.110***	−3.033***
	(0.246)	(0.301)	(0.288)	(0.377)
Observations	2160	2160	2160	2160
Log Likelihood	−376.833	−377.197	−377.252	−376.720
Akaike Inf. Crit.	759.666	760.394	760.503	763.439

* $p < .1$.
** $p < .05$.
*** $p < .01$.

Table 8
Errors in the "look away" trials in the anti-saccade task.

	Dependent variable			
	Errors in the "look away" trials			
	(1)	(2)	(3)	(4)
Classic CRT	−0.951**			−1.246*
	(0.453)			(0.636)
6-Item Numeracy		−0.381		0.473
		(0.438)		(0.588)
VWM			−0.085	−0.043
			(0.091)	(0.091)
Constant	−1.845***	−2.089***	−2.092***	−1.882***
	(0.255)	(0.314)	(0.292)	(0.368)
Observations	2160	2160	2160	2160
Log Likelihood	−685.506	−687.256	−687.198	−685.084
Akaike Inf. Crit.	1377.012	1380.512	1380.396	1380.169

* $p < .1$.
** $p < .05$.
*** $p < .01$.

Table 9
False start in the anti-saccade task.

	Dependent variable			
	False start			
	(1)	(2)	(3)	(4)
Classic CRT	−1.247***			−0.913
	(0.458)			(0.647)
6-Item Numeracy		−1.110**		−0.545*
		(0.443)		(0.605)
VWM			−0.008	0.057
			(0.098)	(0.094)
Constant	−1.421***	−1.330***	−2.057***	−1.398***
	(0.262)	(0.316)	(0.316)	(0.381)
Observations	5049	5049	5049	5049
Log Likelihood	−1958.417	−1958.865	−1961.904	−1957.832
Akaike Inf. Crit.	3922.834	3923.729	3929.808	3925.664

* $p < .1$.
** $p < .05$.
*** $p < .01$.

Table 10
Matching bias in the visual search task.

	Dependent variable			
	Matching bias			
	(1)	(2)	(3)	(4)
Condition 6	1.148***	1.310***	0.989***	1.469***
	(0.188)	(0.225)	(0.208)	(0.275)
Condition 2	−0.672***	−0.433**	−0.679***	−0.606**
	(0.173)	(0.206)	(0.194)	(0.254)
Classic CRT	0.387			−0.132
	(0.246)			(0.341)
Condition 6: classic CRT	−0.848***			−0.304
	(0.324)			(0.456)
Condition 2: classic CRT	0.343			0.731*
	(0.300)			(0.425)
6-Item Numeracy		0.573**		0.606*
		(0.235)		(0.321)
Condition 6: 6-item Numeracy		−0.904***		−0.652
		(0.312)		(0.431)
Condition 2: 6-item Numeracy		−0.090		−0.619
		(0.288)		(0.401)
VWM			0.080	0.064
			(0.049)	(0.050)
Condition 6: VWM			−0.097	−0.057
			(0.065)	(0.067)
Condition 2: VWM			0.063	0.049
			(0.060)	(0.062)

(*continued on next page*)

Table 10 (*continued*)

	Dependent variable			
	Matching bias			
	(1)	(2)	(3)	(4)
Constant	0.185	0.003	0.150	−0.140
	(0.141)	(0.168)	(0.156)	(0.203)
Observations	4742	4742	4742	4742
Log Likelihood	−3039.823	−3041.708	−3042.764	−3034.922
Akaike Inf. Crit.	6093.645	6097.416	6099.529	6095.844

* $p < .1$.
** $p < .05$.
*** $p < .01$.

References

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences, 108*(25), 10367–10371.

Attali, Y., & Bar-Hillel, M. (2020). The false allure of fast lures. *Judgment & Decision Making, 15*(1).

Bacon, A. M., Handley, S. J., Dennis, I., & Newstead, S. E. (2008). Reasoning strategies: The role of working memory and verbal-spatial ability. *European Journal of Cognitive Psychology, 20*(6), 1065–1086.

Baron, J., Scott, S., Fincher, K., & 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*(3), 265–284.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48.

Blacksmith, N., Yang, Y., Behrend, T. S., & Ruark, G. A. (2019). Assessing the validity of inferences from scores on the cognitive reflection test. *Journal of Behavioral Decision Making, 32*(5), 599–612.

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

Brainard, D. H., & Vision, S. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–436.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research, 33*(2), 261–304.

Campitelli, G., & Gerrans, P. (2014). Does the cognitive reflection test measure cognitive reflection? A mathematical modeling approach. *Memory & Cognition, 42*(3), 434–447.

Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability? *American Economic Review, 100*(3), 1238–1260.

Draheim, C., Tsukahara, J. S., Martin, J. D., Mashburn, C. A., & Engle, R. W. (2020). A toolbox approach to improving the measurement of attention control. *Journal of Experimental Psychology, General*, 242–275.

Egeth, H. E., Leonard, C. J., & Leber, A. B. (2010). Why salience is not enough: Reflections on top-down selection in vision. *Acta Psychologica, 135*(2), 130–132.

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist, 49*(8), 709–724.

Erceg, N., & Bubić, A. (2017). One test, five scoring procedures: Different ways of approaching the cognitive reflection test. *Journal of Cognitive Psychology, 29*(3), 381–392.

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

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16*(1), 143–149.

Evans, J. S. B. (2009). In *How many dual-process theories do we need? One, two, or many? In two minds: Dual processes and beyond* (pp. 33–54).

Everling, S., & Fischer, B. (1998). The antisaccade: A review of basic research and clinical studies. *Neuropsychologia, 36*(9), 885–899.

Finucane, M., & Gullion, C. (2010). Developing a tool for measuring the decision-making competence of older adults. *Psychology and Aging, 25*(2), 271–288.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance, 18*(4), 1030–1044.

Frederick, S. (2005). Cognitive reflection and decision making. *The Journal of Economic Perspectives, 19*(4), 25–42.

Fukuda, K., & Vogel, E. K. (2009). Human variation in overriding attentional capture. *Journal of Neuroscience, 29*(27), 8726–8733.

Graffeo, M., Polonio, L., & Bonini, N. (2015). Individual differences in competent consumer choice: The role of cognitive reflection and numeracy skills. *Frontiers in Psychology, 6*, 1–15.

Haigh, M. (2016). Has the standard cognitive reflection test become a victim of its own success? *Advances in Cognitive Psychology, 12*(3), 145–149.

Hallett, P. E. (1978). Primary and secondary saccades to goals defined by instructions. *Vision Research, 18*(10), 1279–1296.

Jackson, S. A., Kleitman, S., Howie, P., & Stankov, L. (2016). Cognitive abilities, monitoring confidence, and control thresholds explain individual differences in heuristics and biases. *Frontiers in Psychology, 7*, 1559–1569.

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye's movement. In J. B. Long, & A. D. Baddeley (Eds.), *IX. Attention and performance* (pp. 187–203). Hillsdale, NJ: Erlbaum. Erlbaum Associates, 187-203.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. *Heuristics and Biases: The Psychology of Intuitive Judgment, 49*, 81–111.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception, 36*(14), 1–16.

Kramer, M. R., Porfido, C. L., & Mitroff, S. R. (2019). Evaluation of strategies to train visual search performance in professional populations. *Current Opinion in Psychology, 29*, 113–118.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence, 14*(4), 389–433.

Liberali, J. M., Reyna, V. F., Furlan, S., Stein, L. M., & Pardo, S. T. (2012). Individual differences in numeracy and cognitive reflection, with implications for biases and fallacies in probability judgment. *Journal of Behavioral Decision Making, 25*(4), 361–381.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*(6657), 279–281.

Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: From psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences, 17*(8), 391–400.

Luck, S. J., Gaspelin, N., Folk, C. L., Remington, R. W., & Theeuwes, J. (2021). Progress toward resolving the attentional capture debate. *Visual Cognition, 29*(1), 1–21.

Mata, F., Sallum, I., Miranda, D., Bechara, A., & Malloy-Diniz, L. F. (2013). Do general intellectual functioning and socioeconomic status account for performance on the children's gambling task? *Frontiers in Neuroscience, 7*, 68–76.

Munoz, D. P., & Everling, S. (2004). Look away: The anti-saccade task and the voluntary control of eye movement. *Nature Reviews Neuroscience, 5*(3), 218–228.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9*(3), 353–383.

Noori, M. (2016). Cognitive reflection as a predictor of susceptibility to behavioral anomalies. *Judgment and Decision making, 11*(1), 114–120.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 437–442.

Pennycook, G., & Rand, D. G. (2019). Cognitive reflection and the 2016 US presidential election. *Personality and Social Psychology Bulletin, 45*(2), 224–239.

Petrova, D., Garcia-Retamero, R., Catena, A., Cokely, E., Carrasco, A. H., Moreno, A. A., & Hernández, J. A. R. (2017). Numeracy predicts risk of pre-hospital decision delay: A retrospective study of acute coronary syndrome survival. *Annals of Behavioral Medicine, 51*(2), 292–306.

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

Rajsic, J., Taylor, J. E. T., & Pratt, J. (2017a). Out of sight, out of mind: Matching bias underlies confir- matory visual search. *Attention, Perception, & Psychophysics, 79*(2), 498–507.

Rajsic, J., Wilson, D. E., & Pratt, J. (2017b). The price of information: Increased inspection costs reduce the confirmation bias in visual search. *The Quarterly Journal of Experimental Psychology, 71*(4), 832–849.

Rajsic, J., Wilson, D. E., & Pratt, J. (2015). Confirmation bias in visual search. *Journal of Experimental Psychology: Human Perception and Performance, 41*(5), 1353–1364.

Rey-Mermet, A., Gade, M., & Oberauer, K. (2018). Should we stop thinking about inhibition? Searching for individual and age differences in inhibition ability. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*(4), 501–526.

Roberts, R. J., Hager, L. D., & Heron, C. (1994). Prefrontal cognitive processes: Working memory and inhibition in the antisaccade task. *Journal of Experimental Psychology: General, 123*(4), 374–393.

Rouder, J. N., Morey, R. D., Morey, C. C., & Cowan, N. (2011). How to measure working memory capacity in the change detection paradigm. *Psychonomic Bulletin and Review, 18*(2), 324–330.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review, 84*(2), 127–190.

Shipstead, Z., Harrison, T. L., & Engle, R. W. (2015). Working memory capacity and the scope and control of attention. *Attention, Perception, & Psychophysics, 77*(6), 1863–1880.

Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Sciences, 9*(1), 16–20.
Sinayev, A., & Peters, E. (2015). Cognitive reflection vs. calculation in decision making. *Frontiers in Psychology, 6*, 532–542.
Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119*(1), 3–22.
Sobel, K. V., Gerrie, M. P., Poole, B. J., & Kane, M. J. (2007). Individual differences in working memory capacity and visual search: The roles of top-down and bottom-up processing. *Psychonomic Bulletin & Review, 14*(5), 840–845.
Stagnaro, M. N., Ross, R. M., Pennycook, G., & Rand, D. G. (2019). Cross-cultural support for a link between analytic thinking and disbelief in god: Evidence from India and the United Kingdom. *Judgment and Decision making, 14*(02), 179–186.
Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences, 23*(5), 645–665.
Stieger, S., & Reips, U.-D. (2016). A limitation of the cognitive reflection test: Familiarity. *PeerJ, 4*, 2395–2397.
Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643–662.
Stupple, E., Gale, M., & Richmond, C. (2013). Working memory, cognitive miserliness and logic as predictors of performance on the cognitive reflection test. In *, 35. proceedings of the annual meeting of the cognitive science society* (pp. 1396–1401).
Szaszi, B., Szollosi, A., Palfi, B., & Aczel, B. (2017). The cognitive reflection test revisited: Exploring the ways individuals solve the test. *Thinking & Reasoning, 23*(3), 207–234.
Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics, 51*(6), 599–606.
Theeuwes, J., Reimann, B., & Mortier, K. (2006). Visual search for featural singletons: No top-down modulation, only bottom-up priming. *Visual Cognition, 14*(4–8), 466–489.
Thomson, K. S., & Oppenheimer, D. M. (2016). Investigating an alternate form of the cognitive reflection test. *Judgment and Decision making, 11*(1), 99–113.
Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The cognitive reflection test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition, 39*(7), 1275–1289.
Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the cognitive reflection test. *Thinking & Reasoning, 20* (2), 147–168.
Unsworth, N., Schrock, J. C., & Engle, R. W. (2004). Working memory capacity and the antisaccade task: Individual differences in voluntary saccade control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(6), 1302–1321.
Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014). Working memory and fluid intelligence: Capacity, attention control, and secondary memory retrieval. *Cognitive Psychology, 71*, 1–26.
Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature, 438*(7067), 500–503.
Walenchok, S. C., Goldinger, S. D., & Hout, M. C. (2020). The confirmation and prevalence biases in visual search reflect separate underlying processes. *Journal of Experimental Psychology: Human Perception and Performance, 46*(3), 274–291.
Weiss, A., Dorrough, A. R., & Schmitz, L. (2021). Analytic atheism in a low-religiosity culture: Examining the relationship between analytic thinking and religious belief in Germany. *Personality and Individual Differences, 178*, 110854–110888.
Weller, J., Dieckmann, N., Tusler, M., C., M. K., Burns, W., & Peters, E. (2013). Development and testing of an abbreviated numeracy scale: A Rasch analysis approach. *Journal of Behavioral Decision Making, 6*(2), 198–212.
Welsh, M., Burns, N., & Delfabbro, P. (2013). The cognitive reflection test: How much more than numerical ability?. In *, 35. Proceedings of the Annual Meeting of the Cognitive Science society* (pp. 1587–1592).