

Individual differences in competent consumer choice: the role of Cognitive Reflection and Numeracy skills

Michele Graffeo, Luca Polonio and Nicolao Bonini

Journal Name:	Frontiers in Psychology
ISSN:	1664-1078
Article type:	Original Research Article
First received on:	20 Feb 2015
Revised on:	21 Apr 2015
Frontiers website link:	www.frontiersin.org

Running head: COGNITIVE REFLECTION AND NUMERACY SKILLS IN CONSUMER CHOICE

Individual differences in competent consumer choice: the role of Cognitive Reflection and Numeracy skills

Michele Graffeo^{1,*}, Luca Polonio², Nicolao Bonini³.

Author Note

¹ Michele Graffeo, Department of Information Engineering and Computer Science, University of Trento; ² Luca Polonio, Center for Mind/Brain Sciences, University of Trento; ³ Nicolao Bonini, Department of Economics and Management, University of Trento.

* Correspondence concerning this article should be addressed to Michele Graffeo, Department of Information Engineering and Computer Science, University of Trento, Via Sommarive 9, 38123 Povo, Italy. E-mail: michele.graffeo@unitn.it.

Abstract

In this paper, we investigate whether cognitive reflection and numeracy skills affect the quality of the consumers' decision-making process in a purchase decision context. In a first (field) experiment, an identical product was on sale in two shops with different initial prices and discounts. One of the two deals was better than the other and the consumers were asked to choose the best one and to describe which arithmetic operations they used to solve the problem; then they were asked to complete the numeracy scale (Lipkus et al., 2001). The choice procedures used by the consumers were classified as "*complete decision approach*" when all the arithmetic operations needed to solve the problem were computed, and as "*partial decision approach*" when only some operations were computed. A mediation model shows that higher numeracy is associated with use of the *complete decision approach*. In turn, this approach is positively associated with the quality of the purchase decision. Given that these findings highlight the importance of the decision processes, in a second (laboratory) experiment we used a supplementary method to study the type of information search used by the participants: eye-tracking. In this experiment the participants were presented with decision problems similar to those used in experiment 1 and they completed the Lipkus numeracy scale and the Cognitive Reflection Test (CRT), (Frederick, 2005). Participants with a high CRT score chose the best deal more frequently, and showed a more profound and detailed information search pattern compared to participants with a low CRT score. Overall, results indicate that higher levels of cognitive reflection and numeracy skills predict the use of a more thorough decision process (measured with two different techniques: retrospective verbal reports and eye movements). In both experiments the decision process is a crucial factor which greatly affects the quality of the purchase decision.

Keywords: Numeracy, CRT, eye movements, decision-making processes, economic choice.

Individual differences in competent consumer choice: the role of Cognitive Reflection and Numeracy skills

Individual differences in cognitive abilities influence important decisions in many domains. Economic decisions, which strongly affect the well-being of the persons, are a primary example. In this domain, Burks, Carpenter, Goette and Rustichini (2009) show that in a large sample of workers (1,000 trainee truckers) those with higher cognitive skills (Numeracy, ability to plan and nonverbal IQ) were more likely to successfully complete their training in a condition where the failure to complete the course involved refunding the training cost. It has also been shown that there is a correlation between higher levels of financial literacy (measured with questions based on the ability to manipulate numbers) and two relevant economic decisions: save for retirement through a private pension plan (Fornero and Monticone, 2011; Lusardi and Mitchell, 2011) and holding stocks (Van Rooij, Lusardi and Alessie, 2011).

A cognitive skill – Numeracy – has a great influence on decisions in the economic context, where often the characteristics of the options are expressed in numerical terms. Data from surveys related to macroeconomic variables that were run in the US, UK, Australia and Canada indicate that higher Numeracy is associated with higher earnings and higher probability to be employed (Charette and Meng, 1998; Chiswick, Lee, and Miller, 2003; McArdle, Smith, and Willis, 2009; McIntosh and Vignoles, 2001). At a microeconomic level, Chen and Rao (2007), show how people often mistakenly evaluate the consequences of a sequence of percentage changes and how their intention to purchase a product is affected by these erroneous calculations. Economic decisions can be problematic in particular for lowly numerate consumers, who resort to a variety of coping strategies (e.g. to buy only known brands) in order to overcome the difficulties presented by the quantitative description of goods (Adkins and Ozanne 2005; Viswanathan, Rosa, and Harris 2005). These findings, in addition to highlighting the central influence of individual differences in cognitive skills, suggest that additional research is needed to understand in greater detail how specific

cognitive skills affect economic decision making. The present study focuses on two cognitive skills: Numeracy and Cognitive Reflection. We investigate how different levels in these two cognitive skills affect the decision process and the quality of choice. The role of Numeracy is investigated in experiment 1. The role of Numeracy and Cognitive Reflection are jointly investigated in Experiment 2. The quality of choice is investigated by using scenarios where the participants choose one of two alternative offers of the same product, with one option that is more valuable than the other. The use of problems that have an optimal solution allows us to investigate the role of cognitive skills on the quality of choice, rather than focusing on choice consistency as in the more familiar literature on framing, inter-temporal choices and similar decision tasks. For a discussion on the differences between problems that involve optimal choices or consistent choices see Reyna and Brainerd (2008).

“Numeracy” has been defined as the ability to comprehend, use and attach meaning to numbers (Peters, Västfjäll, Slovic, Mertz, Mazzocco and Dickert, 2006; Reyna and Brainerd 2008; Reyna et al. 2009). In particular, Numeracy includes the ability to compare magnitudes, and comprehend ratio concepts (including fractions, proportions, percentages, and probabilities). Several studies tried to determine which are the psychological processes underlying the ability to correctly comprehend and manipulate numeric information. Different viewpoints have been suggested: for example Lipkus and Peters (2009) propose a model based on dual system theories (see Kahneman, 2003; Kahneman & Frederick, 2002; Sloman, 1996; Stanovich & West, 2000) where the two systems are responsible for different steps of the comprehension and manipulation of numbers. System 1 is activated during the initial step of this process: when people perceive a numeric stimulus they intuitively evaluate its magnitude (Peters, Slovic, Västfjäll and Mertz, 2008). A subsequent step – number manipulation – is based on logic operations which are part of System 2. Finally, these operations (generated by Systems 1 and 2) concur to the formation of a meaning/interpretation of the numeric information. In this model the functions of the two systems are strictly intertwined and

this is coherent with the idea that good choices are most likely to emerge when the two modes work in concert (Damasio, 1994). Fuzzy-trace theory (see Reyna, 2004; Reyna & Brainerd, 1995) also proposes a dual system approach, which includes verbatim processing (based on precise, detailed or quantitative representations) and gist processing (based on fuzzy, gist or qualitative representations). According to this theory some biases (e.g. ratio bias, denominator neglect) result from gist-based processes: people focus their attention on what they believe to be the more salient piece of information (e.g. the numerator in a fraction or a probability estimate) at the expense of a more complete description of a problem. However, the fuzzy-trace theory considers the gist-based processes as more advanced, compared to the verbatim processes, because they focus on the essential elements of a problems (see Reyna & Brainerd, 1995; Reyna & Lloyd, 2006). Finally, the frequency hypothesis (Gigerenzer, 1994; Gigerenzer & Hoffrage, 1995) posits that humans have directly experienced frequencies throughout their evolutionary history, making frequencies easier to understand compared to artificial constructions such as decimals or probabilities, which are expressed on a normalized scale that does not occur in nature. However, Chapman and Liu (2009) report an experiment where the benefits of the natural frequency format occurred primarily for participants who were high in Numeracy. This result suggests that in the general population the correct comprehension of arithmetic problems can not be assured only by the use of the natural frequencies. For a comparison of the three approaches mentioned above see Reyna and Brainerd (2008). Finally, Del Missier, Mäntylä and Bruine De Bruin (2012) describe Numeracy as a general cognitive ability, which has a sort of intermediate level between basic abilities (the executive function, i.e. control processes involved in the regulation of cognition, see Miyake et al., 2000; Rabbitt, 1997; Royall et al., 2002; Salthouse, 2005) and more complex skills (e.g. the capacity to correctly apply a series of decision rules that involve computations). In particular, a more developed supervision ability - which include the monitoring and revision of working memory contents and the capacity to appropriately inhibit irrelevant information and responses - is associated with higher

Numeracy scores (see also Mäntylä et al., 2007 for a description of this executive function).

The effect of Numeracy on the evaluations and choices has been investigated in different fields such as risk perception, risk communication, pro-social and health-related decisions (see: Dickert, Kleber, Peters, and Slovic, 2011; Greene and Peters, 2009; Hibbard et al., 2007; Lipkus and Peters, 2009; Lipkus, Samsa, and Rimer, 2001; Nelson et al., 2008; Peters, Dieckmann, Dixon, Hibbard, and Mertz, 2007; Peters, Hart, and Fraenkel, 2011; Peters, Hibbard, Slovic, and Dieckmann, 2007). The findings of these studies show, for example, that the economic decisions of low numerate people (e.g. the willingness to donate for charity) are more influenced by changes in numeric presentation format than those of people with higher numeracy (Dickert, Kleber, Peters, and Slovic, 2011). Greene and Peters (2009) found that some contracts, heavily based on numeric information (e.g. health insurance plans) are difficult to understand especially for low numerate people. Consequentially, Peters, Dieckmann et al., (2007) found that strategies which reduce the cognitive load are particularly helpful for low numeracy consumers. Cokely and Kelly (2009) show that there is relationship between cognitive abilities (Numeracy and Reflectiveness, measured through the CRT), decision processes and number of choice that maximize the expected value in a series of gambles. Overall, these studies suggest that highly numerate people draw a more precise and coherent mental representation of numeric information, compared to low numerate people. In addition, highly numerate individuals are more skilled in the integration of numeric information and their choices are less influenced by the way the numeric information is framed, as if they can create a single mental representation of the numeric information, common to different presentation formats. On the contrary, low numerate people tend to focus their attention on some pieces of information, without an adequate integration process (Peters, Slovic, Västfjäll and Mertz, 2008; Peters et al., 2006; Slovic, Monahan and MacGregor, 2000).

Experiment 1

The goal of Experiment 1 is to study the relationship between Numeracy and the decision process

used to choose between different offers of the same product, and how the decision approach affects the quality of choice. Imagine a consumer who considers two alternative deals on the same product. Shop A offers a mobile phone at 180 € with a 15% discount. Shop B offers the same mobile phone at 220 € with a 25% discount. There are no transaction costs (e.g., the two shops are at the same distance from home).

From the Numeracy literature reported in the Introduction, we expect that highly numerate people make a computational effort to select the best deal. For example, in order to find the cheapest option, they calculate the final price of both deals. In more general terms they use a specific decision process that includes performing all the arithmetic operations needed to have a complete description of all the alternatives. We refer to this decision process as a “complete decision approach” to the purchase decision task. We expect that the use of a complete approach to the purchase decision task is positively associated with the selection of the best deal.

On the contrary, we expect that it is less likely that consumers with lower Numeracy scores compute the final prices of the two deals. This might be due to several factors. For example, the math computation is particularly tiring for people with low Numeracy or they don't trust their numerical skills. Regardless of the actual determinant, this also implies a specific decision process in which consumers do not try to perform all arithmetic operations needed to calculate the two final prices (e.g., they compute only one final price and then give up). Alternatively, they simply do not perform any computation at all (e.g., they compare the initial prices or percentage discounts). We refer to this decision process as an “incomplete decision approach” to the purchase decision task. We expect that the use of an incomplete decision approach is positively associated with the selection of the worst deal.

Method

One hundred and fifty-three consumers participated in this study and 42% of the participants were males. In order to increase the ecological validity of our study, we collected the data at the exit of a

shopping mall and asked the participants to choose one of two alternative offers. Three decision scenarios were presented with E-Prime 1.0 (Psychology Software Tools, Inc) software (Schneider, Eschman, & Zuccolotto, 2002a, b), and the participants were randomly assigned to one of the three scenarios, creating three groups of equal size. Each scenario uses a different product (Mobile phone, DVD player, Microwave oven) and the findings show that the choices are not affected by the kind of product. For this reason, in the next section the analyses do not include the factor “product”. Each scenario describes two shops that offer a deal on the same product: one shop offers a low percentage discount on a low initial price (best deal), while the other shop offers a high percentage discount on a high initial price (worst deal) – note that the final prices were not shown to the participants. It may be possible that people associate higher prices with better quality but since we specified that exactly the same product was sold in the two shops it is unlikely that this association occurred during the experiment. The participants decided which deal they preferred (we reversed the order of presentation of the two offers for half of the participants, but there is no order effect). Table 1 reports prices, discounts and final prices.

To investigate the decision process underlying the consumer choice we asked for retrospective verbal reports, a technique widely used in the cognitive science (Ericsson & Simon, 1980) and in the J/DM literature as well (for the use of similar procedures, see Cokely and Kelley, 2009; Payne 1976; Rettinger and Hastie, 2003). The rationale for the use of retrospective verbal protocols in our study is that when people consciously and deliberately consider information (e.g., the computation of the absolute value; the subtraction of the absolute value from the initial price), these processes should be observable in participants’ protocols (Evans, 2008; Sloman, 1996). First we reminded the consumers of the numerical values of the deals (e.g. “Shop A offers a 25% discount on a mobile phone that costs 220 €”), then we asked “When you decided in which shop to buy, did you calculate the corresponding final price (i.e. 220 € minus 25%)?” (Yes vs. No). The question was repeated for the second deal on the same product.

The verbalization task allows us to measure how many consumers computed zero, one or two final prices before making the choice. In order to choose the best deal the consumers should create a complete description of all the possible alternatives, which in this scenario means to calculate two final prices (complete decision approach). Consumers who calculated one or zero final prices had a partial/incomplete description of the options and their choices were most likely based on an approximate estimation of the final prices (partial decision approach).

Finally, the consumers completed the Lipkus et al. (2001) questionnaire. This questionnaire is commonly used in the literature to measure people's Numeracy skill as discussed in the Introduction. In the following analyses we use the number of correct answers as a Numeracy measure (the scale ranges from 0 to 11, the mean score is 7.39, SD: 2.3, median: 8). It is important to note that many of the questions for the Lipkus Numeracy scale ask participants to transform numeric information from the percentage format to the frequency format or vice-versa. The solution of these questions require operations that strongly resemble the operations needed to calculate the absolute value of a discount and the corresponding discounted price, i.e. the operations needed to find the best purchase decision. For this reason we believe that the Lipkus scale is particularly appropriate when measuring the arithmetic skills used in our experiment. In addition, we measured the correlation between the time needed to complete the Lipkus scale and the number of correct answers – the result is not significant ($r(153) = -0.12$, $p = 0.14$). This result indicates that low Numeracy consumers did not try to finish a potentially distressing task as fast as possible, so it is unlikely that the low Numeracy score is caused by a lack of effort or motivation.

INSERT TABLE 1 ABOUT HERE

Results and discussion

In order to analyze the relationships between Numeracy (continuous I.V.), decision approach

(binary Mediator; complete vs. partial decision approach) and choice (binary D.V.; best vs. worst deal), we used a mediation model following the regression approach proposed by Baron and Kenny (1986). Given that the dependent variable is dichotomous, we used a series of binary logistic regressions, as described in Hayes (2009, 2012) and Preacher and Hayes (2004, 2008). The mediation model hypothesizes an effect of the independent variable on the dependent variable through a third explanatory variable, the mediator. The independent variable affects the mediator variable, which in turn affects the dependent variable. In this way the mediator clarifies the nature of the relationship between the independent and dependent variables (MacKinnon, 2008). The mediation model is described in figure 1 and in table 2 we report the descriptive statistics of the independent and dependent variables.

Path C: consumers with low Numeracy chose the worst deal (higher final price) more frequently than the highly numerate consumers (Wald(1) = 5.25; β = 2.28; p = 0.022)¹.

Path A: consumers with low Numeracy used the partial decision approach more often than the highly numerate consumers (Wald(1) = 6.76; β = 2.59; p = 0.009).

Path B indicates a mediation effect: when we include the variable “decision approach”, which has a significant effect on choice (Wald(1) = 4.74; β = 2.18; p = 0.030), Numeracy does not have a significant effect (Wald(1) = 3.57; β = 1.90; p = 0.059). The bootstrapping analysis confirms the mediation effect (indirect effect: 95% CI [.002, .138], based on 5000 bootstrap samples). This result is coherent with our prediction.

INSERT FIGURE 1 ABOUT HERE

INSERT TABLE 2 ABOUT HERE

¹ In both experiments we report standardized beta coefficients, calculated following the procedure proposed by Goodman (1972) and discussed by Menard (2004).

The analyses show that consumers with high Numeracy scores chose the best commercial offer more frequently. At the same time, they were more likely to choose the offer after having calculated both final prices, that is by the means of a complete decision approach to the purchase decision task. On the contrary, consumers with low Numeracy often chose by means of a partial decision approach. They reported that they made the choice without doing any calculation at all, or by the computation of just one final price. This indicates that they chose without knowing the real economic values of the alternatives.

The next step of the mediation analysis shows that consumers who used a complete decision approach chose the best discount more frequently, compared to the consumers who used a partial decision approach (25% vs. 9.4%, $\chi^2(1) = 6.7$; $p = .01$). The final step of the mediation model indicates that the decision approach is a more relevant predictor of purchase decision skill than Numeracy. In other words, consumers with higher Numeracy scores are more likely to use the complete decision approach than consumers with low Numeracy, but if they use the partial decision approach their choices are not different from the choices of the consumers with low Numeracy. In fact, if we consider only the consumers who used a partial approach, a logistic regression analysis indicates that Numeracy does not predict choice quality (Wald(1) = 0.01; $\beta = -0.1$; $p = 0.92$). On the contrary, if we consider only the consumers who used a complete approach, Numeracy significantly predicts choice quality (Wald(1) = 7.40; $\beta = 2.72$; $p = 0.007$). As a final analysis, we compared how much time the consumers needed to decide with the partial or the complete approach and a t-test indicates that consumers who use the partial approach are significantly faster than those who use the complete approach ($M = 48.5s$ vs. $M = 60.2s$; t-test (151) = -2.8 $p = .006$). A similar result has been shown by Cokely and Kelley (2009). The mediation model suggests that even highly numerate consumers might use a partial decision approach to solve the economic decision problem, and this cannot be due to a lack of numeric competence. Said differently, the use of a partial decision

approach to solve the economic problem might depend on plural factors, rather than just Numeracy. To our view, one relevant factor is the attitude about how to approach a decision problem and we used the CRT to explore a specific element of this general attitude: the propensity to act reflexively when facing a decision problem. The role of Cognitive Reflection is fully investigated in the next experiment as a potential factor, and compared to Numeracy as a predictor of the quality of the purchase decision.

Experiment 2

In Experiment 2 we investigate the role of reflectiveness on the decision process and the quality of the purchase decision. Reflectiveness is a broad psychological notion. It has been extensively studied in different fields of psychology (e.g. cognitive, clinical, social and consumer psychology) and economics (e.g. judgment and decision making). While researchers in psychology have focused more on the ability to control impulsive behavior (Eysenck and Eysenck 1978; Puri 1996; Rook and Fisher 1995; see also Grecucci, Giorgetta, Rattin, Guerreschi, Sanfey and Bonini 2014 on the role of impulsivity in compulsive gambling), economists and decision researchers were more focused in hyperbolic discounting, or more generally in the ability to delay gratifications (Monterosso and Ainslie 1999). Studies in psychology have emphasized the distinction between cognitive processes that occur spontaneously and require a moderate level of attention and processes requiring effort, motivation, concentration and the execution of complex learned rules. According to Evans (2008) and Stanovich and West (2000) these different cognitive processes would be supported by two different systems: a basic system (System 1) of effortless and rapid perception that drive intuitive and elementary operations and a more sophisticated system (System 2) which is slow, analytical, and rational.

Following this theoretical framework, the Cognitive Reflection Test (CRT) was conceived as a measure of the ability to control a high accessible (incorrect) answer in favour of a less accessible (correct) answer which requires further deliberation. In the words of the author “CRT measures

‘Cognitive Reflection’ - the ability or disposition to resist reporting the response that first comes to mind” (Frederick, 2005, p. 35). Many researchers since the original publication have used and still use the CRT as a measure of Cognitive Reflection (see for example, Cokely and Kelley, 2009; Fernbach, 2013). In particular, we use the term “cognitive impulsivity” to describe the cognitive style of people with low CRT scores.

The CRT consists of three mathematical word problems² that elicit spontaneous, intuitive and incorrect answers. For example, in one of the CRT problems participants face the following question: “A bat and a ball cost \$1.10. The bat costs \$1.00 more than the ball. How much does the ball cost?” The intuitive answer to this problem is “10 cents”, whereas the correct one is “5 cents”. Frederick’s study shows that CRT is associated to the performance in several decision tasks. For example, in the inter-temporal decision task it has been found that participants with high CRT are more patient, e.g. they chose more frequently the larger reward at a later time, instead of a smaller immediate reward (e.g. they prefer 140\$ in a year to 100\$ now). In risk choice tasks, it has been found that high CRT participants are more willing to take a gamble that has a high expected payoff (instead of a sure gain), compared to low CRT participants (e.g. they choose a 90% to gain 500\$ over a sure gain of 100\$). Subsequent studies report results which are in line with Frederick's findings: for example, Oechssler, Roider and Schmitz (2009) show that people with high scores in the CRT are much less subject to the conjunction fallacy in Linda's problem and they update probability estimates more correctly, compared to low CRT people. According to Campitelli and Labollita (2010), high CRT people are better (compared to low CRT people) at using their background knowledge in order to guess the right answer to question on topics about which they

² About the conceptual distinction between CRT and numeracy it is important to note that, while they share a connection, CRT measures other elements in addition to the ability to manipulate numbers. As shown by Frederick (2005), the correlation between CRT and the SAT Mathematics scores is moderate (.46). More importantly, the CRT has a strong correlation with the preference for patient options in intertemporal choice and with the preference for gambles with the higher expected value (see table 3a, page 34 of the original article). On the contrary, SAT Mathematics scores has little or no correlation with these preferences. In Del Missier, Mäntylä & Bruine de Bruin (2012), numeracy mediates the effect of the executive function “Monitoring and Inhibition”- which is the main determinant behind the answers given to the Cognitive Reflection Test. Overall, the results indicate that the CRT captures other elements of the cognitive processes, in addition to the ability to manipulate numbers.

know little of.

Nofsinger and Varma (2007) found results similar to those of Frederick interviewing a sample of financial advisors. In the domain of the economic decisions Moritz, Hill and Donohue (2013) show that managers with high CRT scores take more profitable decisions. Mixed evidence has been found about the relation between CRT scores and the answers given to moral dilemmas (Paxton, Ungar & Greene, 2011; Royzman, Landy & Leeman, 2014). In Experiment 1, we found that consumers who use a partial approach make their decisions faster and choose the best offer less often compared to those who use a complete approach. This decision pattern is associated to a low level of numerical skills. However, reported findings suggest that, in addition to Numeracy, cognitive impulsivity might inhibit people from making a full computation of final prices. For example, cognitively impulsive consumers might well focus on accessible information to guess which is the best option (e.g. the product with the lowest initial price), although they have the cognitive resources to make the extensive computation (e.g., high Numeracy). Based on previously presented findings, CRT seems to us a good tool to measure cognitive impulsivity in a commercial context. For example, low CRT consumers might rely on accessible information (e.g. lowest initial price; higher percentage discount) to quickly select what looks to them the best deal. This might be the first thought that comes to mind to an impulsive consumer. If not corrected, it might lead the consumer to select the suboptimal offer. This impulsive consumer choice should be reflected also in the decision process. For example, we also expect that low CRT consumers use a simplified information search pattern, compared to high CRT consumers. To that end, in Experiment 2 we record eye movements while participants select one of two offers in 12 discount choice problems. Our objective is to test whether high cognitive reflective consumers are more likely, compared to low cognitive reflective consumers, to use an information search pattern which enable them to compute the final prices, and whether the probability of selecting the best deal depends on the application of this pattern.

A correct procedure to compute the final prices requires the decision makers to compare the initial

price of each option with the associated discount and engage themselves in a mental calculation. In terms of information search pattern, this procedure requires to use vertical eye movements (saccades between an initial price and the relative discount; see figure 2) and long fixations. Vertical eye movements are expected because the initial price and the relative discount are displayed one above the other and this implies that the participants attention should move between these two pieces of information while calculating the final price. Long fixations are expected because the mental calculation of the final price is a complex cognitive operation that is composed of a set of distinct functional processes and longer fixations are commonly associated to deeper processing, such as deliberate consideration of information and planning (Velichkovsky, 1999; Velichkovsky et al., 2002; Glöckner and Herbold 2011; Eivazi and Bednarik, 2011). On the other hand, people who use a partial approach in the selection of the commercial offer might exhibit a different information search pattern. For example, they might compare initial prices and discounts separately. In case that one of the alternatives yields a higher discount with a lower initial price, the decision makers choose that option. Alternatively, they may select the option with the higher discount or with the lower initial price. This simple comparison procedure involves horizontal eye movements (between option transitions) and do not require long fixations because it is less demanding in terms of cognitive effort.

Method

Forty-four persons participated in the experiment, but 9 of them were excluded (five due to poor calibration of the eye-tracking system and excessive number of trial losses, one because of the unusually high mean number of fixations across problems, and three because they reported, after the end of the experiment, that they were previously aware of some of the answers of the CRT). The mean age of the remaining 35 participants (14 males and 21 females) is 24.5 (SD = 3.47). The experiment was performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki, and all participants gave informed consent prior to admission into the

study.

After entering the lab, participants were informed of the nature of the “purchase decision ” task. Participants responded to 12 discount choice problems presented in a random order. For each problem the participants decided whether to buy a product (e.g. DVD player) at shop A or B. The two shops, A and B, sold the same identical product which was on sale with different initial prices and percentage discounts. For half of the problems the best deal is the one with the lowest initial price (see figure 2-A), for the other half it is the one with the highest initial price (figure 2-B). For each discount choice problem, participants had to select their response by pressing the corresponding key on the keyboard. Before starting the experiment, participants performed 4 practice trials. Choices and fixation times were automatically registered. After the discount choice experiment, subjects were asked to complete the Lipkus et al. Numeracy scale (Lipkus, Samsa, & Rimer, 2001) and the Cognitive Reflection Test (Frederick, 2005).

INSERT FIGURE 2 ABOUT HERE

Eye-tracking procedures: Participants were seated in a chair with a soft head restraint to ensure a viewing distance of 60 cm to the monitor. Presentation of the stimuli was performed using a custom made program written using the Matlab Psychophysical toolbox. Eye movements were monitored and recorded using an Eyelink II system (SR. Research Ontario Canada) with a sampling rate of 500 Hz. A fixation was defined as an interval in which gaze was focused within 1° of visual angle for at least 100 ms. (Manor and Gordon 2003). A nine-point calibration was performed at the beginning of the experiment. Calibration phase was repeated until the difference between the positions of the points on the screen and the corresponding eye locations was less than 1°. After the calibration phase, a nine-point validation phase was performed (similarly to the calibration phase) to make sure that the calibration was accurate. Recalibrations were performed if needed, and eye-

tracking was stopped if these were unsuccessful. Before the beginning of each trial a drift correction was performed. Then, a fixation point was presented in the same position as the last point of the drift correction for 300 milliseconds. The fixation point was located at the bottom of the two possible options, outside the area covered by the picture representing the discount choice problem. Each discount choice problem was presented after the fixation point and remained on the screen until a response was made. Eye movements were recorded during the purchase decision problem display (figure 3).

INSERT FIGURE 3 ABOUT HERE

Eye-tracking-data

For each of the 12 problems, we defined 4 areas of interest (AOIs), as shown in figure 4. None of the AOIs were adjacent to the others. This allowed us to avoid the possibility that small errors in the calibration procedures could result in an incorrect attribution of the eye movements parameters (fixations and saccades, see below). This is especially useful for those situations in which the parameter was located on the border of an AOI. AOIs do not cover the picture representing the discount choice problem entirely (but less than 8% of the screen) and never overlap. In this way, AOIs include only fixations and saccades whose interpretation is not ambiguous. Those parameters that were not located inside the AOIs were not considered in the analysis. However, although a large part of the picture is not included in any AOI, the majority of fixations observed fell inside the AOIs (88%). For each trial, we recorded the *average length of fixations* for each of the 4 AOIs, the *number of times* that the participant looked within an AOI, and the number and type of *saccades*, which correspond to the transitions made by the eye from one AOI to the next.

INSERT FIGURE 4 ABOUT HERE

Results

Choices: Numeracy and CRT

First, we investigated the influence of Numeracy on the participants' ability to select the best offer. As in Experiment 1, Numeracy is measured by the number of correct answers given to the Lipkus scale and we treated Numeracy as a continuous variable (scores range from 1 to 11). The binary dependent variable is the quality of commercial choice (i.e. the selection of the best deal). We decided to use a mixed regression model because the participants answered twelve discount choice problems which are considerably different from one another. They involve different products and different price ranges. Also, the absolute difference and the ratio of the initial and final prices change. This variability can be explained using the concept of “random factors”, i.e., variables whose levels are drawn “from a large (potentially very large) population in which the individuals differ in many ways, but we do not know exactly how or why they differ” (Crawley, 2007; p. 628). The mixed regression model (unlike ANOVA and simple linear regression) provides a way to take into account the added variation caused by differences between the levels of the random factors (Baayen, Davidson, & Bates, 2008; Fugard, Pfeifer, Mayerhofer, & Kleiter, 2011). As a consequence we included in our model a random factor called “Problem”, that discriminates among the twelve discount choice problems. Results from the logistic mixed model show the importance of Numeracy: higher Numeracy significantly increases the choice of the best options ($\beta = 0.47$, $Z = 2.23$, $p = 0.03$). In a second model, in addition to Numeracy, we added the factor CRT, considering the four levels of ability (from 0 to 3 correct answers). This model shows that Numeracy does not longer affect the choice of the best option ($\beta = 0.26$, $Z = 1.18$, $p = .32$), while CRT does ($\beta = 0.82$, $Z = 3.42$, $p < .001$). Given that these variables explore two related constructs we decided to verify the level of collinearity between them, so we calculated the Variance Inflation Factor (VIF). Usually the largest VIF is taken to be a measure of the seriousness of collinearity among the predictors and the

higher the value, the higher the collinearity. The definition of “high” is arbitrary but values higher than 4 or 5 are commonly considered an index of excessive correlation among explanatory variables. We found an index of 1.07 which allows us to exclude collinearity problems between the variables. As a second test, we also calculated the correlation between CRT and Numeracy to check for possible differences between our study and previous studies. We opted for the Spearman correlation coefficient, as it is not possible to assume neither that the variables are normally distributed nor that there is a linear correlation between them (both of which are necessary assumptions for the use of the Pearson correlation coefficient). Results show that the variables (CRT and Numeracy) are marginally correlated (Spearman correlation, $r(35) = 0.29$, $p = .09$). In previous studies, Cokely and Kelly (2009) reported a significant correlation of 0.31. A stronger correlation was found in Finucane and Gullion (2010; $r = .53$) and in Liberali et al. (2012; $r = 0.51$ and $r = 0.40$). Despite the small size of our sample due to the constraints imposed by the eye tracking device (participants have to be tested individually), the degree of correlation we found between Numeracy and CRT is in line with that obtained by Cokely and Kelly (2009).

Results from the mixed model regression analyses are congruent with those obtained in experiment 1, and they highlight the role of Cognitive Reflection in the selection of the best commercial offer. In figure 5 we reported the relationship between performance on the CRT and the proportion of best offers chosen by the participants. Table 3 details the summary statistics for the questionnaires and the eye-tracking analysis, while table 4 describe the parameters of the mixed models presented below.

INSERT FIGURE 5 ABOUT HERE

INSERT TABLE 3 ABOUT HERE

INSERT TABLE 4 ABOUT HERE

CRT and Eye movements

As shown in the previous section, the effect of Numeracy scores on the proportion of best offers chosen is no longer significant when included in the same model with the CRT scores; therefore, in the following analyses, we focus only on the relationship between CRT and eye movements. Our research hypothesis predicts that participants with different levels of CRT use different decision processes and we investigate this prediction using three eye movement parameters (number and length of fixations and the Search Index calculated on the vertical/horizontal saccades).

Analysis of Saccades. Vertical saccades are movements from one of the top cells (prices) to its corresponding lower cell (discount) or vice-versa (within-option or cue-based search). Horizontal saccades are movements from left to right (or vice-versa), so they connect the two prices or the two discounts (between-option or option-based search). In order to summarize the participants' search patterns we calculated the Search Index (SI; see Norman and Schulte-Mecklenbeck, 2010 for a theoretical description and Glöckner and Betsch, 2008 for an application of the SI to an eye-tracking study). The SI is calculated by subtracting the number of between-option transitions from the number of within-option transitions and dividing this difference by the sum of both transitions. SI ranges from -1 to +1 and negative values indicate more between-option searches, while positive values indicate more within-option searches. Within-option transitions are necessary to link the initial price with the corresponding discount, in order to calculate the final price. Conversely, between options transitions in our purchase decision task do not give rise to any economically relevant comparison.

We run a linear mixed model to test the hypothesis that participants who obtained high scores in the

Cognitive Reflection Test were more prone to use within-option transitions. Results indicate that higher CRT scores significantly increase the proportion of vertical (within-option) transitions (higher SI, $\beta = 0.27$, $t = 5.68$, $p < .001$).

Analysis of the average length of fixations. The data were treated in the following way: we calculated the average length of fixations for each AOI and used the natural logarithm transformation of the values in order to approximate a normal distribution (Newell & Rosenbloom, 1981). The average length of fixations for the two price areas are highly correlated, ($r(35) = .94$, $p < .001$), like the average length of fixations for the two discount areas ($r(35) = .70$, $p < .001$). Conversely, the correlations between the average length of fixations for the each of the two prices with each of the two discounts are not significant. Given these results, we considered the areas of the two prices as a single area of interest and, similarly, we treated the two discounts as a single area of interest. The factors length of fixations for the Price Area (LF-P) and length of fixations for the Discount Area (LF-D) indicate the average length of fixations on the two macro-areas.

We run a linear mixed model to evaluate the hypothesis that the score obtained in the Cognitive Reflection Test is positively correlated with LF-P. Results indicate that a higher score in the CRT significantly increases the LF-P ($\beta = 0.33$, $t = 7.21$, $p < .001$). The same analysis was repeated to evaluate the association between CRT and LF-D, but this time the predictor was not statistically significant ($\beta = 0.08$, $t = 1.57$, $p = .12$).

Fixation Counts. We recorded how many times the participants looked at a certain AOI and, paralleling the analyses on the average fixation length, we aggregated the AOIs into two macro-areas: the Price Area and the Discount Area. The factors number of fixations in the Price (NF-P) and Discount Areas (NF-D) indicate the average number of fixations within each of the two macro-areas. Then, we ran two linear mixed models and tested whether the score obtained in the CRT is positively correlated with the NF-P and NF-D factors. Similarly to what we observed in the analysis of the average length of fixations, we found that participants who have a higher score in the CRT

have a higher number of fixations in the Price Area ($\beta = 0.20$, $t = 4.15$, $p < .001$) but not in the Discount Area ($\beta = 0.06$, $t = 1.16$, $p = .25$).

Overall, results of the saccades and fixation analyses show that participants with different CRT scores used different information search patterns. In particular, the performance on the CRT is positively related to the proportion of within-option transitions and the average length and number of fixations in the Price Area.

Relationship between the depth of the elaboration process and the within-option search strategy.

We studied the relationship between fixations (average length and number of fixations) and saccades (Search Index) because we expected to find an association between the depth of the elaboration process and the within-option search strategy. For the SI a linear mixed model with the factors LF-P and NF-P shows that the average length and number of fixations is significantly associated with the proportion of within-option transitions. More precisely, the proportion of within-option transitions increases when the length ($\beta = 0.27$, $t = 5.55$, $p < .001$) and the number of fixations ($\beta = 0.23$, $t = 5.04$, $p < .001$) increase within the Price Area. This result indicates that the participants who processed the problem information more deeply used a within-option evaluation strategy.

Eye movements and Choice Accuracy

Finally, we test the relationship between the information search patterns and the probability of selecting the best offer. We run a series of logistic mixed model that includes the eye movement parameters that had a significant effect in the previous analyses (SI, the LF-P, and the NF-P) using the backward elimination procedure and we test whether they are good predictors of choice accuracy (i.e. how often the participants chose the more advantageous option). The initial model (AIC = 497.2) which includes the three predictors shows that LF-P is marginally significant ($\beta = 0.49$, $Z = 1.77$, $p = 0.07$), while NF-P ($\beta = 0.12$, $Z = 0.80$, $p = 0.426$) and SI ($\beta = -0.08$, $Z = -0.44$, $p = 0.663$) are not. We eliminated SI, the least relevant predictor, and we run a model with the

variables LF-P and NF-P. The fit of the model improves (AIC = 495.4), LF-P has a marginally significant effect ($\beta = 0.49$, $Z = 1.72$, $p = 0.08$), while NF-P does not ($\beta = 0.14$, $Z = 0.71$, $p = 0.48$). Finally, we eliminated the variable NF-P and now the model shows another improvement in the goodness-of-fit (AIC = 493.9) and a fully significant effect of LF-P ($\beta = 0.53$, $Z = 2.27$, $p = 0.023$).

Information search patterns for high and low CRT scores

In order to give a general picture of the information search patterns used by participants with different CRT scores we divided them into two groups - high and low CRT - and we compared their search patterns. Participants who gave two or three correct answers were classified as “high CRT” (17 participants) and those who gave one or zero correct answers as “low CRT” (18 participants). As shown in figure 6 the average length of fixations on the price areas for high CRT participants is almost 100 milliseconds (ms) longer compared to low CRT participants (low CRT = 334 ms; high CRT = 423 ms). Conversely no differences were found on the Discount areas (low CRT = 249 ms; high CRT = 258 ms). These results suggest that high CRT participants are more likely to engage themselves in the calculation of the final prices and that the mental calculation of the final discounted prices was performed while fixating on the price areas, which greatly simplified the maintenance of the initial prices in the working memory. The two groups differ also in the number of vertical saccades (low CRT = 6.05; high CRT = 7.5) but not in the number of horizontal saccades (low CRT = 3.55; high CRT = 3.35), which is in accordance with the fact that only saccades between price and discount values are useful for the mental calculation.

In general, we suggest that high CRT participants are more likely to remain focused on the initial prices, using vertical saccades to read and maintain the discount values in working memory. This information search pattern is suitable for the mental calculation of the final price. Conversely, low CRT participants are more likely to use a less specific information search pattern which does not include the mental calculation of the final price (see figure 6 below).

INSERT FIGURE 6 ABOUT HERE

Conclusions

Reported findings suggest that the ability to choose the best option in a simple purchase decision problem depends not only on numerical skills, but also on cognitive impulsivity as measured by the Cognitive Reflection Test. We also found that the CRT predicts the information search pattern used which, in turn, affects the quality of purchase decision. Specifically, we found that consumers with low levels of Cognitive Reflection do not use a sufficiently detailed information search pattern and pick a low proportion of the more advantageous offers.

General discussion

Consumers must often decide which is the most advantageous economic option among a list of alternatives, and the literature shows that sometimes they choose poorly in conditions of risk or uncertainty (Bonini, Tentori and Osherson, 2004; Kahneman, 2003; Shu, 2007; Starmer, 2000).

Consumers also show poor performance when coping with riskless choice. For example, when they have to find which is the cheapest goods bundle for a certain product or – as in the present study – which is the best deal (see Armstrong, 2010 for a comprehensive analysis on this point).

Our research aims to expand the current knowledge about the characteristics of economic behavior by studying the influence of two cognitive skills: Numeracy and Cognitive Reflection. While it has already been shown that higher levels of cognitive skills (e.g. ability to plan, memory, etc.) are associated with better choices (Burks, Carpenter, Goette and Rustichini, 2009; Christelis, Jappelli and Padula, 2010), the comprehension of the link between decision processes and cognitive skills is still limited. In addition, the association between CRT scores and choices has been little studied in the context of the economic domain - one of the few examples is the study by Finucane and Gullion (2010).

In a series of purchase decision problems, our experimental findings indicate that consumers with

high CRT scores and high Numeracy are significantly more likely to make the more advantageous choice, compared to people who are more cognitively impulsive and/or less numerate. Starting from our experimental hypotheses this result can be explained by the different decision processes that characterize participants with high/low scores in the CRT and Numeracy Test. In experiment 1 we show that consumers with low Numeracy use a partial decision approach, i.e. they use only part of the numeric information provided and they do not compute all the operations needed to identify the best option. Hence, they select the worst option much more frequently, compared to highly numerate consumers. In experiment 2 we found that both Cognitive Reflection and Numeracy have an influence, with a particularly strong effect of former variable on the aptness of the purchase decisions. At the same time, the analyses on the eye movement data show that more cognitively impulsive people adopt a shallow search pattern, which is associated with worse performance in the purchase decision task. On the whole, both experiments support our experimental hypotheses that people with greater cognitive skills are more likely to use thorough decision-making processes, which produce a higher number of optimal choices.

We believe that the reported findings can contribute to the growing literature on the vulnerability of consumers (see OECD 2010 consumer policy toolkit). As shown by a recent survey on proficiency in key information-processing skills among working-age adults (OECD 2013), insufficient Numeracy is a widespread phenomenon: only one adult out of three is able to act upon and correctly interpret relatively simple data and statistics in texts, tables and graphs. The great extent of the phenomenon indicates that Numeracy is certainly one factor in consumer vulnerability, as shown also in our study. This vulnerability appears also in very common economic problems, where consumers make decisions based upon the computation of simple numeric information such as integers or percentages (see Yin and Dubinsky, 2004 on the influence of comparative price formats and Graffeo and Bonini, 2015, on the relationship between Numeracy and comparative price formats). However, reported findings add another factor to the vulnerability list: lack of Cognitive

Reflection or cognitive impulsivity. Consumers with low CRT scores tend to choose poorly, most likely because their first impression of “the best deal” is not corrected by further information processing (e.g., the computation of the final prices).

These findings are of interest for both marketing experts and consumer policy specialists since there is a recent and growing concern in the European Union about unfair commercial practices, and consumer rights in general (European Parliament and Council 2005; 2011). Furthermore, findings suggest that the attributes that characterize the effective decision-maker include particular cognitive skills other than financial or mathematical knowledge. This conclusion is congruent with the broad definition of “financial capability” provided by OECD (PISA 2012, comma 43, page 13) where it is stated that: “Additionally [to knowledge and understanding of financial concepts and risks], financial literacy involves skill in managing the emotional and psychological factors that influence financial decision making”. We believe that the ability to counteract the first (and false) impression of goodness in a commercial setting is one of these crucial psychological factors.

Finally, our findings show that the use of some specific decision processes is an indicator that certain factors of vulnerability are at work; a more detailed knowledge of these processes could facilitate the development of comprehensive guidelines that illustrate how to prevent some recurrent mistakes. In this way it would be possible to mitigate the influence of the vulnerability factors, in particular within socially relevant domains.

Acknowledgments:

The authors contributed in equal measure to the development of the paper. Michele Graffeo gratefully acknowledges the financial support of the European Union and of the Provincia Autonoma di Trento (7th Framework Program “People”, Marie Curie Action – COFUND, project PEOCC). The authors wish to thank M. Dellantonio for their help in the data collection of experiment 1 and V. Corelli for experiment 2.

References.

- Adkins, N. R., & Ozanne, J. L. (2005). The low literate consumer. *Journal of Consumer Research*, 32(1), 93-105.
- Armstrong, J. S. (2010). *Persuasive Advertising: Evidence-based Principles*. London: Palgrave Macmillan.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412.
- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology*, 51(6), 1173.
- Bonini, N., Tentori, K. & Osherson, D. (2004). “A Different Conjunction Fallacy”, *Mind and Language*, 19(2), 199-210.
- Burks, S. V., Carpenter, J. P., Goette, L., & Rustichini, A. (2009). Cognitive skills affect economic preferences, strategic behavior, and job attachment. *Proceedings of the National Academy of Sciences*, 106(19), 7745-7750.
- Campitelli, G., & Labollita, M. (2010). Correlations of Cognitive Reflection with judgments and choices. *Judgment and Decision Making*, 5(3), 182-191.
- Charette, M. F., & Meng, R. (1998). The determinants of literacy and Numeracy, and the effect of literacy and Numeracy on labour market outcomes. *Canadian Journal of Economics*, 31(3), 495-517.
- Chapman, G. B., & Liu, J. (2009). Numeracy, frequency, and Bayesian reasoning. *Judgment and Decision Making*, 4(1), 34-40.
- Chen, H. A., & Rao, A. R. (2007). When two plus two is not equal to four: Errors in processing multiple percentage changes. *Journal of Consumer Research*, 34(3), 327-340.
- Chiswick, B. R., Lee, Y. L., & Miller, P. W. (2003). Schooling, literacy, Numeracy and labour

market success. *Economic Record*, 79(245), 165-181.

Christelis, D., Jappelli, T., & Padula, M. (2010). Cognitive abilities and portfolio choice. *European Economic Review*, 54(1), 18-38.

Cokely, E. T. & Kelley, C. M. (2009). Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation. *Judgment and Decision Making*, 4(1), 20-33.

Crawley, M. J. (2007). *The R Book*. John Wiley, New York.

Damasio, A. (1994). *Descartes' error: Emotion, reason, and the human brain*. New York: Avon.

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

Dickert, S., Kleber, J., Peters, E., & Slovic, P. (2011). Numeracy as a precursor to pro-social behavior: The impact of Numeracy and presentation format on the cognitive mechanisms underlying donation decisions. *Judgment and Decision Making*, 6(7), 638-650.

Eivazi, S., & Bednarik, R. (2011). Predicting problem-solving behavior and performance levels from visual attention data. In *Proc. Workshop on Eye Gaze in Intelligent Human Machine Interaction at IUI* (pp. 9-16).

Eysenck, S. B., & Eysenck, H. J. (1978). Impulsiveness and venturesomeness: Their position in a dimensional system of personality description. *Psychological reports*, 43(3f), 1247-1255.

European Parliament and Council. (2005). *Council Directive 2005/29/EC of 11 May 2005 on unfair business to consumer commercial practices in the internal market and amending Council Directive 84/450/EEC, Directives 97/7/EC, 98/27/EC and 2002/65/EC of the European Parliament and of the Council and Regulation (EC) No. 2006/2004 of the European Parliament and of the Council*. Luxembourg: European Council.

European Parliament and Council. (2011). *Council Directive 2011/83/EU of 25 October 2011 on*

consumer rights, amending Council Directive 93/13/EEC and Directive 1999/44/EC of the European Parliament and of the Council and repealing Council Directive 85/577/EEC and Directive 97/7/EC of the European Parliament and of the Council.

- Evans, J. S. B. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual of Review Psychology, 59*, 255-278.
- Fernbach, P. M., Sloman, S. A., Louis, R. S., & Shube, J. N. (2013). Explanation fiends and foes: How mechanistic detail determines understanding and preference. *Journal of Consumer Research, 39*(5), 1115-1131.
- Finucane, M. L., & Gullion, C. M. (2010). Developing a tool for measuring the decision-making competence of older adults. *Psychology and aging, 25*(2), 271-288.
- Fornero, E., & Monticone, C. (2011). Financial literacy and pension plan participation in Italy. *Journal of Pension Economics and Finance, 10*(04), 547-564.
- Frederick, S. (2005). Cognitive Reflection and decision making. *Journal of Economic perspectives, 19*(4), 25-42.
- Fugard, A. J. B., Pfeifer, N., Mayerhofer, B., & Kleiter, G. D. (2011). How people interpret conditionals: Shifts toward the conditional event. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(3), 635–648.
- Gigerenzer, G. (1994). Why the distinction between single-event probabilities and frequencies is important for psychology (and vice versa). In G. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 129–161). New York: Wiley.
- Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological review, 102*(4), 684.
- Glöckner, A., & Betsch, T. (2008). Multiple-reason decision making based on automatic processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(5), 1055–1075.
- Glöckner, A., & Herbold, A. K. (2011). An eye-tracking study on information processing in risky

decisions: Evidence for compensatory strategies based on automatic processes. *Journal of Behavioral Decision Making*, 24(1), 71-98.

Goodman, L. A. (1972). A modified multiple regression approach to the analysis of dichotomous variables. *American Sociological Review*, 37, 28-46.

Graffeo, M. & Bonini, N. (2015). On the evaluation of savings: the role of Numeracy. *Neuropsychological Trends*, 17, 31-35. doi: 10.7358/neur-2015-017-graf

Grecucci A, Giorgetta C, Rattin A, Guerreschi C, Sanfey AG & Bonini, N. (2014) Time Devours Things: How Impulsivity and Time Affect Temporal Decisions in Pathological Gamblers. *PLoS ONE* 9(10), 8 October 2014, Article number e109197. doi:10.1371/journal.pone.0109197

Greene, J., & Peters, E. (2009). Medicaid consumers and informed decision making. *Health care financing review*, 30(3), 25.

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication monographs*, 76(4), 408-420.

Hayes, A. F. (2012). PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling [White paper]. Retrieved from <http://www.afhayes.com/public/process2012.pdf>

Hibbard, J. H., Peters, E., Dixon, A., & Tusler, M. (2007). Consumer Competencies and the Use of Comparative Quality Information It Isn't Just about Literacy. *Medical Care Research and Review*, 64(4), 379-394.

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American economic review*, 93(5), 1449-1475.

Kahneman, D., & Frederick, S. (2002). Representativeness Revisited: Attribute Substitution in Intuitive Judgment. In T. Gilovich, D. Griffin & D. Kahneman (Eds.), *Heuristics and Biases: The Psychology of Intuitive Judgment* (pp. 49–81). New York: Cambridge

University Press.

- 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.
- Lipkus, I. M., & Peters, E. (2009). Understanding the role of Numeracy in health: proposed theoretical framework and practical insights. *Health Education & Behavior*, 36(6), 1065-1081.
- Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a Numeracy scale among highly educated samples. *Medical Decision Making*, 21(1), 37-44.
- Lusardi, A., & Mitchell, O. S. (2011). Financial literacy around the world: an overview. *Journal of Pension Economics and Finance*, 10(04), 497-508.
- MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. Routledge.
- Manor, B. R., & Gordon, E. (2003). Defining the temporal threshold for ocular fixation in freeviewing visuocognitive tasks. *Journal of Neuroscience Methods*, 128(1-2), 85-93.
- Mäntylä, T., Carelli, M. G., & Forman, H. (2007). Time monitoring and executive functioning in children and adults. *Journal of Experimental Child Psychology*, 96(1), 1-19.
- McArdle, J. J., Smith, J. P., & Willis, R. (2009). *Cognition and economic outcomes in the Health and Retirement Survey* (Working Paper No. 15266). Retrieved from National Bureau of Economic Research website: <http://www.nber.org/papers/w15266>.
- McIntosh, S., & Vignoles, A. (2001). Measuring and assessing the impact of basic skills on labour market outcomes. *Oxford Economic Papers*, 53(3), 453-481.
- Menard, S. (2004). Six approaches to calculating standardized logistic regression coefficients. *The American Statistician*, 58(3), 218-223.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal

lobe” tasks: A latent variable analysis. *Cognitive psychology*, 41(1), 49-100.

Monterosso, J., & Ainslie, G. (1999). Beyond discounting: possible experimental models of impulse control. *Psychopharmacology*, 146(4), 339-347.

Moritz, B. B., Hill, A. V., & Donohue, K. L. (2013). Individual differences in the newsvendor problem: Behavior and cognitive reflection. *Journal of Operations Management*, 31(1), 72-85.

Nelson, W., Reyna, V. F., Fagerlin, A., Lipkus, I., & Peters, E. (2008). Clinical implications of Numeracy: theory and practice. *Annals of behavioral medicine*, 35(3), 261-274.

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J.R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1-55). Lawrence Erlbaum Associates.

Norman, E., & Schulte-Mecklenbeck, M. (2010). Take a quick click at that! Mouselab and eye-tracking as tools to measure intuition . In Glöckner, A., & Witteman, C. (Eds.), *Foundations for tracing intuition : challenges and methods* (pp. 24 - 44). New York : Psychology Press.

Nofsinger, J. R., & Varma, A. (2007). How analytical is your financial advisor? *Financial Services Review*, 16(4), 245-260.

Oechssler, J., Roider, A., & Schmitz, P. W. (2009). Cognitive abilities and behavioral biases. *Journal of Economic Behavior & Organization*, 72(1), 147-152.

OECD (2010). *Consumer Policy Toolkit*, OECD Publishing, doi: <http://dx.doi.org/10.1787/9789264079663-en>.

OECD (2013). *OECD Skills Outlook 2013: First Results from the Survey of Adult Skills*, OECD Publishing.

Paxton, J. M., Ungar, L., & Greene, J. D. (2012). Reflection and reasoning in moral judgment. *Cognitive Science*, 36(1), 163-177.

Payne, J.W. (1976). Task complexity and contingent processing in decision-making — information

- search and protocol analysis. *Organizational Behavior and Human Performance*, 16(2), 366–387.
- Peters, E., Dieckmann, N., Dixon, A., Hibbard, J. H., & Mertz, C. K. (2007). Less is more in presenting quality information to consumers. *Medical Care Research and Review*, 64(2), 169-190.
- Peters, E., Hart, P. S., & Fraenkel, L. (2011). Informing Patients The Influence of Numeracy, Framing, and Format of Side Effect Information on Risk Perceptions. *Medical Decision Making*, 31(3), 432-436.
- Peters, E., Hibbard, J., Slovic, P., & Dieckmann, N. (2007). Numeracy skill and the communication, comprehension, and use of risk-benefit information. *Health Affairs*, 26(3), 741-748.
- Peters, E., Slovic, P., Västfjäll, D., & Mertz, C. K. (2008). Intuitive numbers guide decisions. *Judgment and Decision Making*, 3(8), 619-635.
- Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological Science*, 17(5), 407-413
- PISA (2012). FINANCIAL LITERACY ASSESSMENT FRAMEWORK, Doc: FinLit_Frmwrk_PISA12, April 2012
- Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior research methods, instruments, & computers*, 36(4), 717-731.
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior research methods*, 40(3), 879-891.
- Puri, M. (1996). Commercial banks in investment banking conflict of interest or certification role?. *Journal of Financial Economics*, 40(3), 373-401.
- Rabbitt, P. (1997). Introduction: Methodologies and models in the study of executive function.

Methodology of frontal and executive function, 1-38.

- Rettinger, D. A., & Hastie, R. (2003). Comprehension and decision making. In S. L. Schneider & J. Shanteau (Eds.), *Emerging perspectives on judgment and decision research*, pp. 165–201. New York, NY: Cambridge University Press.
- Reyna, V. F. (2004). How People Make Decisions That Involve Risk A Dual-Processes Approach. *Current Directions in Psychological Science*, *13*(2), 60-66.
- Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences*, *7*(1), 1-75.
- Reyna, V. F., & Brainerd, C. J. (2008). Numeracy, ratio bias, and denominator neglect in judgments of risk and probability. *Learning and Individual Differences*, *18*(1), 89-107.
- Reyna, V. F., & Lloyd, F. J. (2006). Physician decision-making and cardiac risk: Effects of knowledge, risk perception, risk tolerance, and fuzzy processing. *Journal of Experimental Psychology: Applied*, *12* (3), 179–195.
- Reyna, V. F., Nelson, W. L., Han, P. K., & Dieckmann, N. F. (2009). How Numeracy influences risk comprehension and medical decision making. *Psychological bulletin*, *135*(6), 943.
- Rook, D. W., & Fisher, R. J. (1995). Normative influences on impulsive buying behavior. *Journal of consumer research*, *22* (3),305-313.
- Royall, D. R., Lauterbach, E. C., Cummings, J. L., Reeve, A., Rummans, T. A., Kaufer, D. I., ... & Coffey, C. E. (2002). Executive control function. A review of its promise and challenges for clinical research. A report from the Committee on Research of the American Neuropsychiatric Association. *The Journal of neuropsychiatry and clinical neurosciences*, *14*(4), 377-405.
- Royzman, E. B., Landy, J. F., & Leeman, R. F. (2015). Are thoughtful people more utilitarian? CRT as a unique predictor of moral minimalism in the dilemmatic context. *Cognitive Science*, *39*, 325-352. DOI: 10.1111/cogs.12136

- Salthouse, T. A. (2005). Relations between cognitive abilities and measures of executive functioning. *Neuropsychology, 19*(4), 532.
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002a). *E-Prime user's guide*. Pittsburgh: Psychology Software Tools Inc.
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002b). *E-Prime reference guide*. Pittsburgh: Psychology Software Tools Inc.
- Shu, Suzanne B., *Choosing for the long run: Making trade-offs in multiperiod borrowing*, Working Paper, UCLA, 2007.
- Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological bulletin, 119*(1), 3-22
- Slovic, P., Monahan, J., & MacGregor, D. G. (2000). Violence risk assessment and risk communication: the effects of using actual cases, providing instruction, and employing probability versus frequency formats. *Law and human behavior, 24*(3), 271.
- Stanovich, K. E. and West, R. F., (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences, 23*(5), 645-665.
- Starmer, C. (2000). Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of economic literature, 38*(2), 332-382.
- Van Rooij, M., Lusardi, A., & Alessie, R. (2011). Financial literacy and stock market participation. *Journal of Financial Economics, 101*(2), 449-472.
- Velichkovsky, B. M. (1999). From levels of processing to stratification of cognition: Converging evidence from three domains of research. In B. H. Challis , & B. M. Velichkovsky (Eds.), *Stratification in cognition and consciousness* (pp. 203–235). Amsterdam, Netherlands: John Benjamins Publishing Company.
- Velichkovsky, B. M., Rothert, A., Kopf, M., Dornhöfer, S. M., & Joos, M. (2002). Towards an express-diagnostics for level of processing and hazard perception. *Transportation Research*

Part F: Traffic Psychology and Behaviour, 5(2), 145-156.

- Viswanathan, M., Rosa, J. A., & Harris, J. E. (2005). Decision making and coping of functionally illiterate consumers and some implications for marketing management. *Journal of Marketing*, 69(1), 15-31.
- Yin, W., & Dubinsky, A. J. (2004). Framing effects of coupon face value on coupon redemption: A literature review with propositions. *Journal of Marketing Management*, 20(7-8), 877-896.

Running head: COGNITIVE REFLECTION AND NUMERACY SKILLS IN CONSUMER CHOICE

Table 1: List of the products for shop A and B, with the corresponding initial prices and discounts, exp. 1

Product	Shop A			Shop B		
	Initial price (€)	Discount (%)	Final price (€)	Initial price (€)	Discount (%)	Final price (€)
Mobile phone	180	15	153	220	25	165
DVD player	140	30	98	180	40	108
Microwave oven	70	20	56	90	30	63

Running head: COGNITIVE REFLECTION AND NUMERACY SKILLS IN CONSUMER CHOICE

Table 2: Descriptive statistics of the independent and dependent variables of exp. 1

	Average	Standard deviation
Numeracy - Number of correct responses	7.39	2.3
Time needed to complete the decision task (ms)	55025	26270
	Proportions	
Decision approach	Complete approach: 55.6%	Partial approach: 44.4%
Choice	Best option: 83.7%	Worst option: 16.3%

Table 3: Descriptive statistics of the independent and dependent variables of exp. 2. Eye-tracking statistics are calculated during the discount choice problem presentation.

	Average	Standard deviation
CRT - Number of correct responses	1.40	1.31
Numeracy - Number of correct responses	8.11	1.94
Response time (ms)	20604	14586
Mean fix duration (ms)	323	83
Mean fix duration (shop A) (ms)	317	83
Mean fix duration (shop B) (ms)	314	75
Mean fix duration (Prices) (ms)	366	111
Mean fix duration (Discounts) (ms)	249	43
Total fixation time per decision (ms)	17061	14503
Total AOI fixation time per decision (ms)	16115	11702
Total fixation time per decision (shop A) (ms)	8874	6935
Total fixation time per decision (shop B) (ms)	7255	5589
Total fixation time per decision (Prices) (ms)	12319	10110
Total fixation time per decision (Discounts) (ms)	3804	2271
Number of total fixations per decision	49	35
Number of AOI fixations per decision	43	24
Number of AOI fixations per decision (Shop A)	23	13
Number of AOI fixations per decision (Shop B)	20	11
Number of between-option transitions per decision (prices)	4.90	1.94
Number of between-option transitions per decision (discounts)	1.99	0.99
Number of within-option transitions per decision (shop A)	7.15	4.10
Number of within-option transitions per decision (shop B)	6.41	3.93

Table 4: A summary of the generalized linear mixed models used to test the relation among accuracy and questionnaires scores (Numeracy and CRT), among eye-movements parameters and CRT scores and among accuracy and eye-movements parameters.

Dependent variables	Predictor variables	Coefficient Estimates	Z/t value	P value
Choices	Questionnaires		Z value	
Choices (0/1)	Numeracy (0/11)	0.47	2.23	0.03
N° of obs = 420	AIC = 494.4	BIC = 506.5	LogL = -244.2	
Choices (0/1)	Numeracy (0/11)	0.26	1.18	P = 0.32
	CRT (0/3)	0.82	3.42	P < 0.001
N° of obs = 420	AIC = 484.5	BIC = 500.7	LogL = -238.3	
Eye-movements parameters	Questionnaires		t value	
SI	CRT (0/3)	0.27	5.68	P < 0.001
N° of obs = 420	AIC = 120.4	BIC = 136.6	LogL = -56.6	
LF-P	CRT (0/3)	0.33	7.21	P < 0.001
N° of obs = 420	AIC = 264.8	BIC = 281.0	LogL = -128.4	
LF-D	CRT (0/3)	0.08	1.57	P = 0.12
N° of obs = 420	AIC = 90.5	BIC = 106.7	LogL = -41.3	
NF-P	CRT (0/3)	0.20	4.15	P < 0.001
N° of obs = 420	AIC = 3805.3	BIC = 3821.5	LogL = -1898.7	
NF-D	CRT (0/3)	0.06	1.16	P = 0.25
N° of obs = 420	AIC = 3115.7	BIC = 3131.9	LogL = -1553.8	
Eye-movements parameter	Eye-movements parameters		t value	
SI	LF-P	0.27	5.55	P < 0.001
	NF-P	0.23	5.04	P < 0.001
N° of obs = 420	AIC = 65.7	BIC = 85.9	LogL = -27.8	
Choices	Eye-movements parameters		Z value	
Choices (0/1)	LF-P	0.49	1.77	P = 0.07
	NF-P	0.12	0.80	P = 0.43
	SI	-0.08	-0.44	P = 0.66
N° of obs = 420	AIC = 497.2	BIC = 517.4	LogL = -243.6	
Choices (0/1)	LF-P	0.49	1.72	P = 0.08
	NF-P	0.14	0.71	P = 0.48
N° of obs = 420	AIC = 495.4	BIC = 511.6	LogL = -243.7	
Choices (0/1)	LF-P	0.53	2.27	P = 0.02
N° of obs = 420	AIC = 493.9	BIC = 506.0	LogL = -244.0	

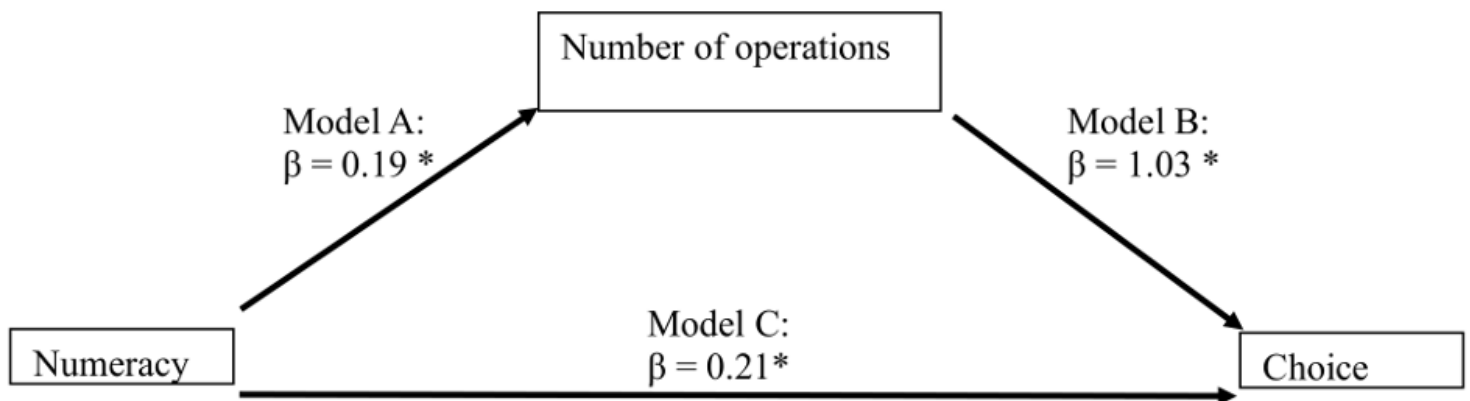


Figure 1: The mediation model for experiment 1 shows that Numeracy predicts the Decision Approach, which in turn predicts Choice. * Significant at $p < .05$.

		DVD player	
		Shop A	Shop B
Initial Price	90 €	70 €	
Discount	30 %	20 %	

		Jeans	
		Shop A	Shop B
Initial Price	90 €	80 €	
Discount	25 %	10 %	

Figure 2: Two examples of economic problems. In the first example (A) the more worthwhile option is the one with the low initial price (Shop A), in the second example (B) the best option is the one with the high initial price (Shop B).

Figure 3.TIF

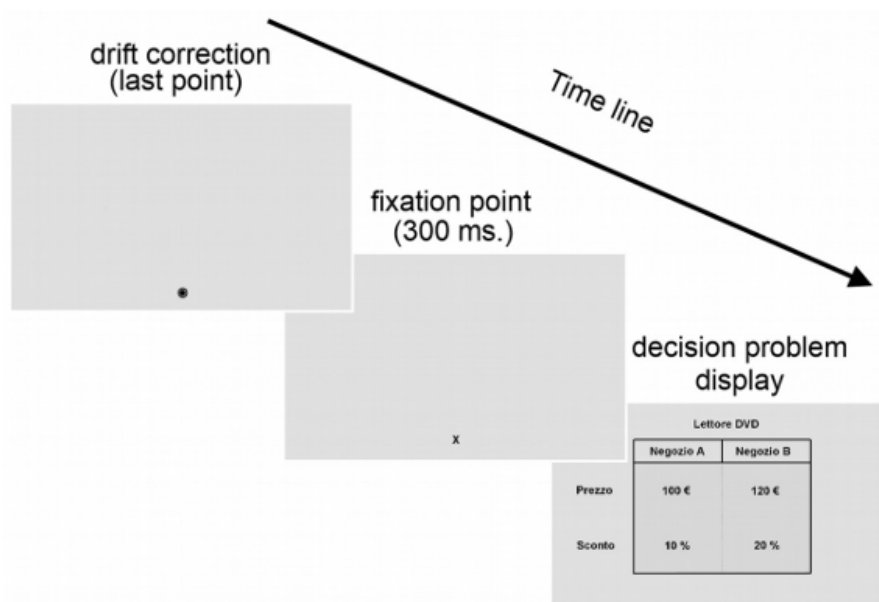


Figure 3: Task and time course of a trial. At the end of the drift correction, a fixation point was displayed for 300 millisecond on the bottom of the two possible options. Then, the economic problem was displayed onscreen until a response was made.

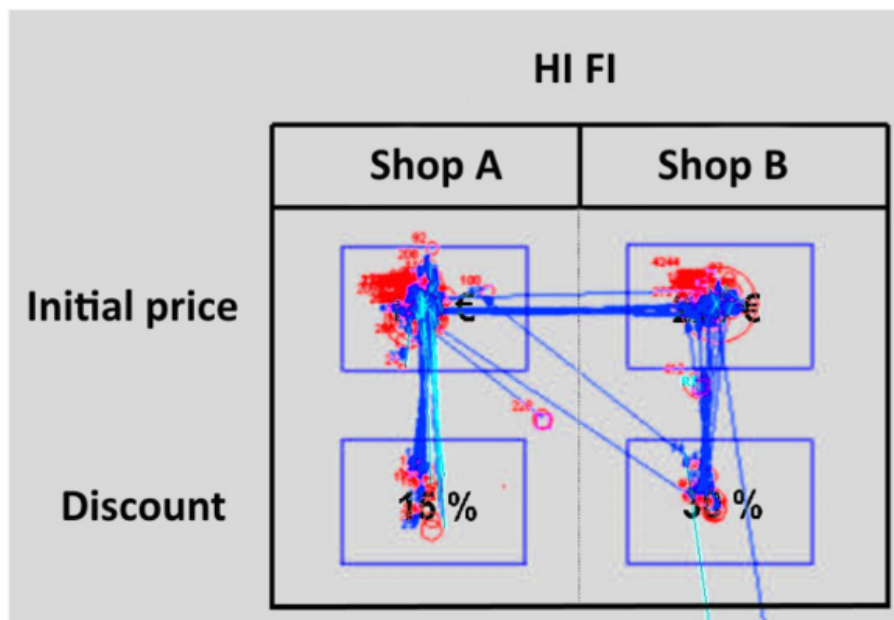


Figure 4: Eye movement for one participant while responding to the discount choice problem. Blue lines indicate saccades and red circles the location of fixations (the size of the circle is proportional to the length of the fixation). The red number next to the red circle indicates the length of the fixation. Blinks are indicated with light blue lines. The blue rectangles (not visible to the participants) represent the AOIs.

Figure 5.TIF

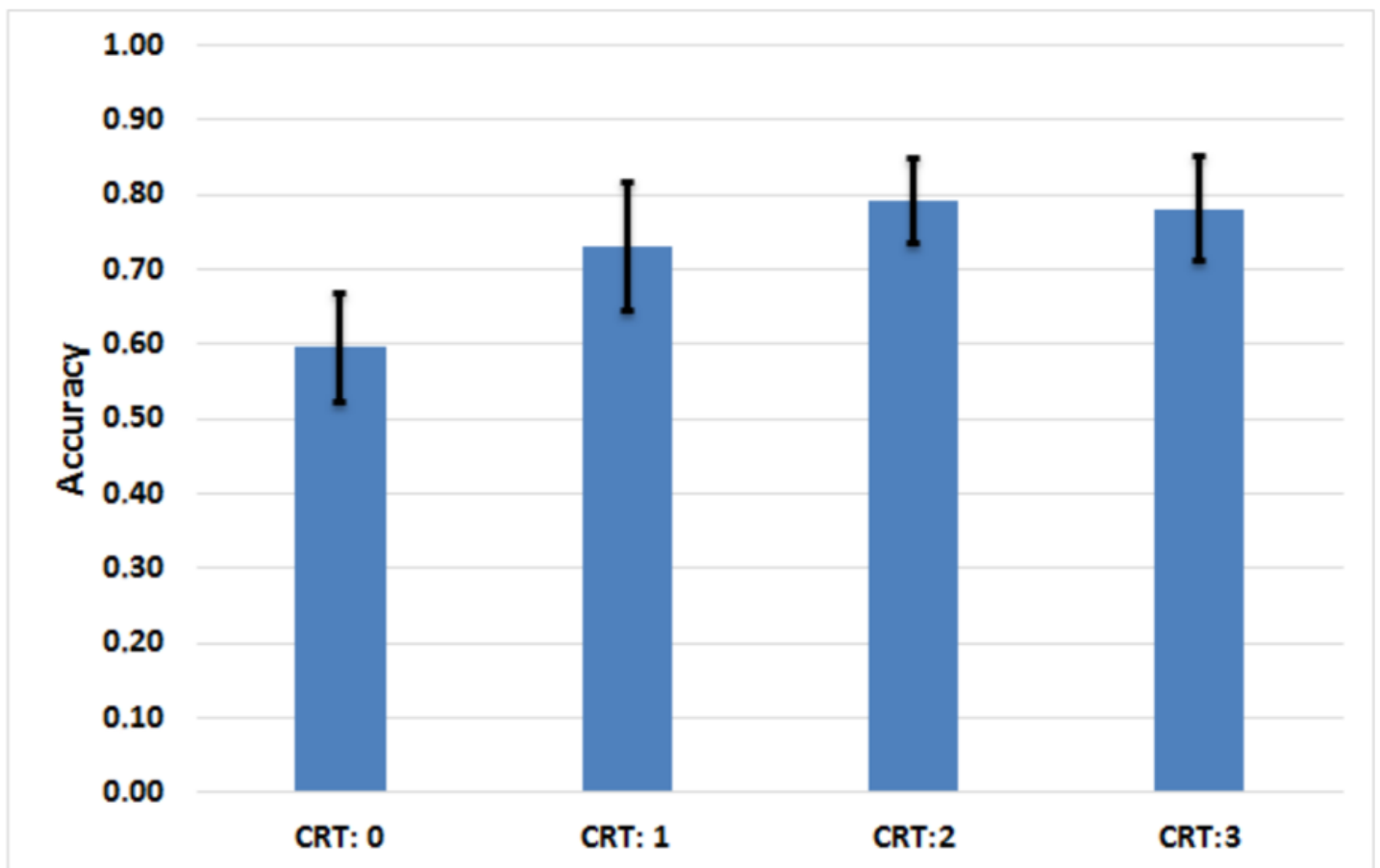


Figure 5: Assessment of the relationship between the performance on the CRT and the level of accuracy (proportion of best offers chosen) achieved by the participants in the 12 discount choice problems (mean and standard deviation).

Figure 6.TIF

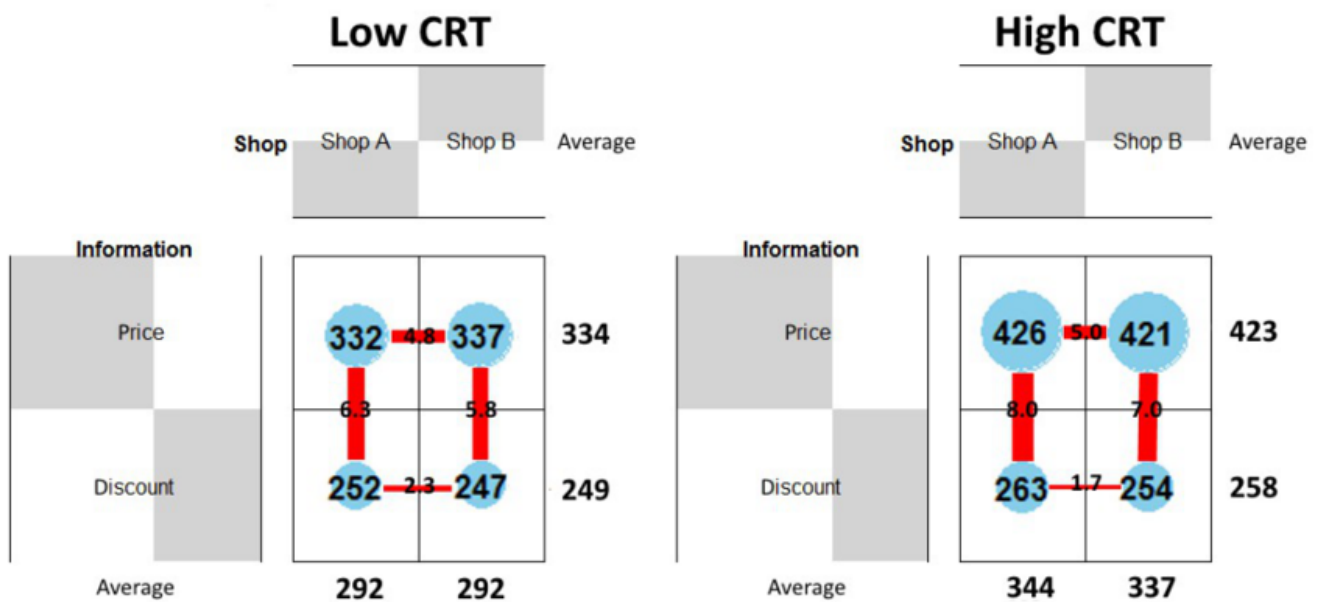


Figure 6: Information processing for low (left) and high (right) CRT participants. The numbers within the blue circles indicates the average length (in milliseconds) of fixations for the four AOIs (the radius of the circles is proportional to the average length of fixations). The numbers superimposed on the red lines indicates the average number of saccades for the four types of transitions (the thickness of the lines is proportional to the average number of saccades).