Fast and wrong: An eye-tracking exploration on how low "cognitive reflection" people analyze and choose commercial packages with multi-dimensional prices.

Alessia Dorigoni*, Nicolao Bonini (presenter)*

*Dipartimento di Economia e Management,
Via Inama, 1 38121 Trento
Università di Trento (Italy)

dorigoni.alessia@gmail.com

nicolao.bonini@unitn.it

720 - Behavioural Finance Issues

Fast and wrong: An eye-tracking exploration on how low "cognitive reflection" people analyze and choose commercial packages with multi-dimensional prices.

Abstract

Fixation durations can be used to provide insights into cognitive processes; longer fixations are associated with an increasing level of processing. Based on the literature on individual differences on cognitive processes, we analyzed gaze behavior during a purchase decision context in order to understand if the levels of cognitive reflection affect the type of price-information processing, and in turn, the quality of choice. The items of the cognitive reflection test (CRT) were designed to trigger an erroneous but intuitive response; scores on this test predict a wide range of judgment and decision biases. The aim of this paper is to investigate the role of the inhibitory component of the CRT, measured by the number of non-intuitive answers of the CRT-ICS (Inhibitory Control Score), on the decision-making process in a purchase context, and its impact on the quality of choice. Participants were presented with two websites selling the same commercial package and asked to choose one. The two alternative packages were displayed by four price dimensions. Fixation durations and the direction of the information search were recorded using eye-tracking technology (Eye Link 1000 Plus). After the final choices were recorded, participants were asked to fill several cognitive reflection tests, and a numeracy test.

We found a worse choice quality for people with low CRT-ICS (e.g., selection of the more expensive package). The underlying cognitive processes were investigated, and two possible explanations for the low-quality choice finding were tested by analyzing gaze behavior. A "semantic grouping of multi-dimensional prices" hypothesis (more impulsive participants will tend to group multi-dimensional prices by their semantic relatedness and neglect the unrelated one). A "superficial price-information processing" hypothesis (more impulsive participants will tend to analyze all displayed multi-dimensional prices more quickly). Results support the second hypothesis and show

that participants with lower cognitive reflection spend less time to look at all displayed price dimensions which, in turn, leads to a worse choice accuracy. The results are interesting because they highlight that cognitive reflection can manifest not only in our thinking but how we allocate attention to the information and the environment.
The exploration of cognitive processes in decision-making, particularly in a purchase context and the role of cognitive reflection in how individuals process price information have a theoretical relevance for understanding how people process price information and make choices. Furthermore, it underscores the importance of attention allocation, as observed through eye-tracking and fixation durations, in understanding these cognitive processes.

Keywords: online purchases, eye tracking, cognitive reflection, categorization process, mental accounting, multidimensional prices

---

1. Introduction

Individual differences in judgment and decision-making have been extensively studied in recent years [1], often investigating whether some individuals are more prone to relying on heuristics and biases in their decisions. Such research has revealed that the presence and magnitude of decision biases can be predicted by an individual's cognitive reflection, which is defined as the tendency to reflect on a question instead of reporting the first, potentially erroneous, response that comes to mind. In the present study, we investigate whether cognitive reflection predicts the tendency to ignore relevant, but seemingly unrelated, price information in purchasing decisions.

Although cognitive reflection refers to a broad tendency to rely on intuitions rather than deliberation in decision-making, it is investigated through using a Cognitive Reflection Test (CRT, [2]). As it was first developed, the CRT requires participants to solve three problems that are designed to automatically generate intuitive responses. Many studies have shown that CRT scores correlate with heuristics and biases [3–5], seemingly because those with lower cognitive reflection base their decisions on System 1 psychological processes that are fast, automatic and effortless [6]. For example, the negative correlation between CRT scores and the number of conjunction fallacies suggests that lower cognitive reflection means increased reliance on similarity-based judgments such as the representativeness heuristic (See [7], [8]). Lower CRT scores are also associated with a more superficial information search pattern and a lack of adequate information integration, further supporting the relatively higher reliance on heuristics over deliberation [9], [10], [11]. Several authors agree that the CRT is not just another numeracy scale [12] and some of them studied the link between CRT and numeracy and their effect on decision-making [13], [14]. For example, Campitelli and Labollita [15] considered the possibility that the effect of cognitive reflection on decision-making performance could be mediated by numeracy. However, the correlations they found were not in line with this hypothesis. Sirota and Juanchich proposed to test whether numeracy can mediate the effect of cognitive reflection on Bayesian reasoning [16] and they found that numeracy does not mediate the effect of cognitive reflection on Bayesian performance. The role of numeracy in performance was clarified and cognitive reflection was identified as a new determinant of Bayesian performance. Even in simple lotteries, superior risky decisions associated with cognitive abilities and controlled cognition can reflect metacognitive dynamics and elaborative heuristic search processes, rather than normative calculations. Other research indicates that individual differences in risky decision-making may also arise from variations in one's general knowledge and understanding of probabilities such as numeracy [17]. Sobkow and colleagues [18] found that multiple numeric competencies predicted decision-making beyond fluid intelligence and cognitive reflection; moreover, a combination of different cognitive abilities contributed to a better understanding of decision outcomes. Also in other studies lower cognitive reflection ultimately leads to lower quality purchase decisions([19], [20], [21]); subjects with higher cognitive ability play closer to the Nash-equilibrium in a p-beauty contest [22]; actual willingness to pay for various consumer goods can be manipulated by an uninformative anchor and the effect is lower in people with higher cognitive reflection [23] and the scale risk aversion and impatience decrease with higher cognitive ability ([24]).

Superficial information-processing of displayed prices. In order to make rational purchasing decisions, people must compute the total price by the means of a mental calculation. For example, when shown a car lease priced as "232€ a month for 36 months" people need to multiply 232 by 36 to get the total price. Mental calculation requires effort and is related to a deep information processing. Because lower cognitive reflection is associated with increased use of System 1 psychological processes, purchasing decisions that allow for heuristic decision-making may be made differently by those lower in cognitive reflection.

A consumer is confronted with a multi-dimensional price when more than one set of numbers is displayed for the given purchase, as in the above example (For a review on multi-dimensional prices see [25]). Other examples of such transactions are "a 125€ jacket on sale for 13€ off" or "a mobile phone for 499€ plus 110€ taxes and 9.90€ for overnight shipping and handling". Multi-dimensional pricing is frequently used in sales, and under these circumstances the consumer needs to compute the net price by integration of all price dimensions to make a rational choice ([26], p. 327).

In purchasing decisions that involve multiple dimensions of pricing, the heuristic information-processing could involve attending to some price dimensions and ignoring others, which is called attribute or dimension neglect (See [27] for a discussion).

Pricing research, for example, has found that people might well overlook a price dimension when confronted with a multi-dimensional price. For example, when shown a car lease priced as "232€ a month for 36 months", people might well focus on the monthly payment ignoring how the total price needs to be computed (e.g., multiplication of 232 by 36). Morwitz and colleagues ([28], Study 2), found that slightly less than 25% of participants completely ignored the surcharge for shipping and handling when a telephone's price was presented in a partitioned fashion (see Abraham and Hamilton [29] for a meta-analysis on partitioned prices). Many factors might account for why a price dimension might be ignored: the complexity of the mental operations required to compute the total or net price, the magnitude of the product price under consideration, the relative importance of the price dimension, the way a discount is communicated (absolute vs. proportional value), presentation order, and the importance of the purchase itself are just a few examples (see [26] for a review on multi-dimensional prices). Also, in the stated preference literature, it has been shown how people in a choice card experiment often do not attend to all displayed choice attributes [27].

In this study, we investigated a potential reason for the price-dimension neglect which, to the best of our knowledge, has not been considered. We argue that price-dimensions that are perceived as semantically unrelated to the object of the purchase will be neglected ("semantic-neglect"), and that this neglect will be more pronounced in those with lower cognitive reflection.

The "semantic neglect" hypothesis being tested in our study is derived from categorization theory ([30], [31], [32]). One principle in categorization theory is that people group together (e.g., categorize) elements that are perceived semantically related (e.g., because of their similarity or because they share a set of properties). For example, a robin but not a bulldog is easily/quickly recognized as a member of the category "BIRD". Furthermore, according to the theory of prototype [30], members of a category have different levels of categorical membership as a function of their similarity with a prototype ("typicality effect"). For example, due to a difference in their typicality, a robin and a canary are more easily/quickly recognized as members of the category "BIRD" than a penguin is.

Henderson and Peterson [33], first tried to extend the psychological processes of categorization to mental representation of economic events by arguing the conceptual equivalence between categorization theory and mental accounting (For a formal categorization-based model of mental accounting see [34]; see [35] on mental accounting theory).

If the principles of categorization (e.g., the principles of grouping and graded membership within the group) are used also for the mental representation of a purchase presented by a multidimensional price then it is plausible to expect that atypical price dimensions in the context of a given purchase will not be considered (e.g., the fee for the use of the credit card in the context of the evaluation of holiday package, financial costs for the purchase of a car that is bought by borrowing money or the administration costs for the purchase of an airline ticket). Put in other words, purchasers should be less likely to attend, and therefore integrate, a price dimension into their subjective valuation of a purchase when the price dimension is semantically unrelated to the purchase.

To investigate whether a semantic-neglect bias exists in the context of multidimensional pricing, and whether it is more pronounced under low cognitive reflection, we asked participants to choose between pairs of commercial packages that were presented by a multidimensional price.

Based on the previous demonstrations of more heuristic-based decision-making in those with lower cognitive reflection [2], [36], [37], we expected that CRT scores would predict how individuals choose between commercial offers, and whether semantic relatedness of price dimensions predicts how much attention they are given in purchasing decisions. Specifically, we expected that those with lower CRT scores would make poorer choices (e.g., the preference for the package with the highest net price) than high CRT participants when semantically unrelated price dimensions were critical to choosing the lower-priced option. In addition to looking at purchasing decisions, we used fixation duration to measure the extent to which different price dimensions were attended during purchasing decisions, considering their association with deeper cognitive processing ([38]; [39]; [40]; [41]; [42]). Analysis of fixation to infer the processing of multi-attribute purchasing decisions has been used successfully in numerous studies ([43]; [44]).

The paper by Kim and Kachersky ([45]) aims to conceptualize dimensions of price salience, which influences price perceptions and deal evaluations, especially when a price consists of more than a single number (multi-dimensional prices). The authors develop a framework based on existing research on price salience and general salience. They identify four basic dimensions of price salience: visual, semantic, computational, and magnitude salience, arguing that each dimension has a unique influence on price perceptions. The aim of our paper is to show that also the perceived semantic relatedness amongst the price dimensions in the context of a given purchase can affect the deal evaluations during the decision-making process.

2. Pilot Study

2.1. Method

The aim of the pilot study was to measure the perceived semantic relatedness amongst the price-dimensions in the context of a given purchase (e.g., a washing machine, a holiday package, etc.). If all price-dimensions of a purchase are perceived semantically related (e.g., because they all are the cost components of the market good), a unique semantic grouping should be found. If people perceive different semantic relatedness amongst the price-dimensions (e.g., because of their differences in typicality), then different semantic groupings should emerge. For example, in the case of a holiday package, if the "price of the room", the "price of the breakfast" and the "credit card fee" are all perceived as semantically related because they are the three cost components of the considered purchase, then one semantic grouping should be found. However, if people perceive the "credit card fee" as semantically unrelated to the other two prices that represent the typical dimensions of a hotel room, then the "credit card fee" should not be grouped together with the "bed and breakfast" costs whereas these last two ones should be frequently grouped together.

It should be noted that the hypothesized "semantic-neglect bias" previously discussed assumes that the several price-dimensions of a purchase might not be equally typical within an activated category of purchase (e.g., the purchase of a hotel room; the purchase of a car; the purchase of an airline ticket). Said differently, there cannot be a "semantic neglect" when all displayed price-dimensions are perceived as equally typical dimensions of the considered purchase. Therefore, it is imperative to first demonstrate that people might well perceive different semantic relatedness between the price dimensions in the context of a given purchase.

3. Participants

Twenty students from a local university participated in the pilot study, and the card sorting task was used. The card-sorting technique is a typical method to investigate the semantic space of a concept ([46] and [47]). It has also been developed for designing mainframe menu systems ([48]) and has been recently used in the user centered design field in order to create information architectures of effective and user-friendly websites [49] and [50]. Thus, we used the card sorting task to control whether the several price dimensions of a purchase are all perceived as semantically related (e.g., only one semantic grouping should emerge) or not.

4. Design

We collected data as suggested in Tullis and Wood [51]. For each commercial package, the participants were presented with six cards that represented its price-dimensions. In table 1 are presented the six cards for each of the three commercial packages.

5. Procedure

Following Tullis and Wood ([52]), participants were asked to group the cards by following a criterion which made sense for them. The names of the price dimensions to be grouped were printed on the cards in an easily readable font. To detect which price dimensions are easily grouped together and which are not, we computed, for each package, how many times each price-dimension was grouped together with all the others.

Holiday by the sea	**Washing machine**	**Gym membership**
room charge	WM cost	GM cost
tourist tax	transport cost	registration card cost
breakfast	mounting cost	discount
credit card fee	extended warranty	sauna
price discount	disposal cost	Locker cost
wi-fi fee	appropriate detergent	doctor’s appointment

Table 1: Multidimensional prices for each commercial package

6. Results

For all three purchases, results show that the displayed price dimensions are not grouped in one single cluster. Instead, different groupings of the initial set of six cards emerge. For example, as shown in Figure 1, “the price of the room” and “the price of the breakfast” are frequently grouped together whereas “the fee for the credit card” is not.

There is no reason to expect such clustering of prices from a rational perspective because all prices should equally be considered to get the total price. However, people differentiate between the price dimensions of the same transaction. These results support the theoretical prediction (based on categorization theory previously discussed) that, although the displayed multi-dimensional prices relate to the same purchase and are presented with the same unit scale (e.g., a euro price), they might [53] well be perceived as differently semantically related between them. Said differently, the displayed price dimensions are posted into different semantic groups (e.g., one price relate to the main decision object and the other prices are more or less related to it). To the best of our knowledge, this is the first evidence of different semantic groupings of the multi-dimensional prices of a purchase.

Findings are reported in Tables A1, A2, A3 (Appendix A). In order to visualize the results, we used Cluster Dendograms (Figure 1, Figure 2, Figure 3) that represent the hierarchical cluster analysis ([54]). We chose as similar dimensions (in green) the three closest ones, and as dissimilar dimension (in red) the one farthest from the similar ones (based also from the frequency tables in the Appendix A; in the first column are reported the four price dimensions for each commercial package used in the next Experiment).

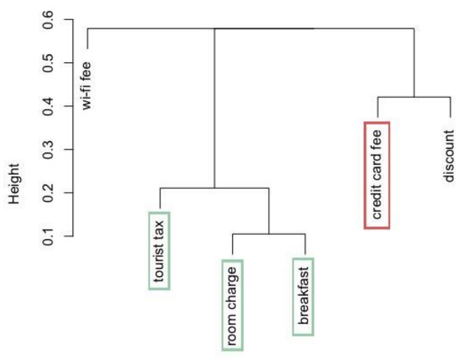


Figure 1: Cluster Dendogram for the "holiday by the sea" task.

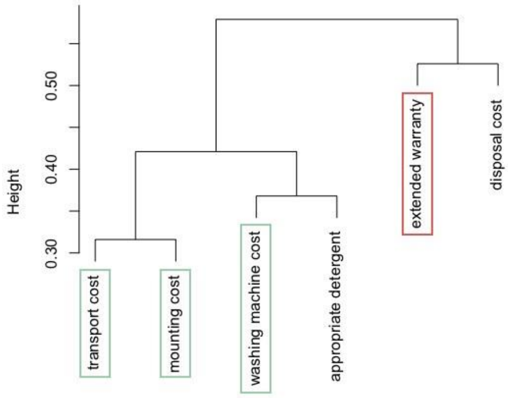


Figure 2: Cluster Dendogram for the "washing machine" task.

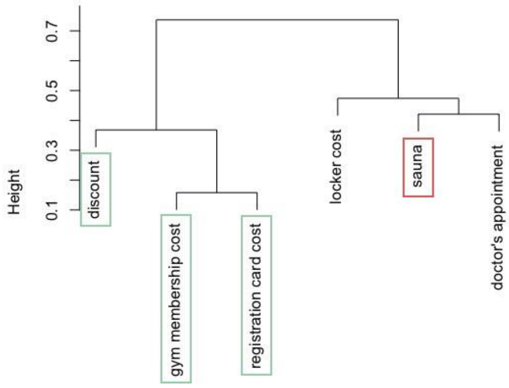


Figure 3: Cluster Dendogram for the "gym membership" task.

7. Discussion

Based on these results, in the next experiment each of three purchases was described by four price dimensions, that is the three similar price dimensions and the fourth dissimilar one. The three similar dimensions were those that were more often grouped together (e.g., they belong to the same semantic group). The dissimilar dimension was that which was less often grouped with the three similar ones (e.g., it does not belong to the same semantic group of the three similar dimensions). In Appendix A, the first column of Table A1, A2 and A3 summarizes the four price dimensions for each purchase used in the next experiment. Let us call the first three dimensions "similar price dimensions" (denoted as "s1", "s2" and "s3"), and the last dimension "dissimilar price dimension" (denoted as "d").

8. Experiment

Having established that the price dimensions in our stimuli differ in their semantic relatedness, we conducted an experiment to test whether semantically dissimilar price dimensions 'd' are neglected, and whether cognitive reflection is associated with this semantic-price neglection in purchasing decisions. To measure cognitive reflection, we combined several cognitive reflection measures and counted the number of intuitive responses given across each. Several authors have proposed a different coding schemes of CRT responses to better measure the inhibition component of cognitive reflection ([55], [56], [57], [58]). For example, Sinayev and Peters ([56]) argued that coding CRT responses into two categories (intuitive responses and non-intuitive responses) allows the separation of cognitive reflection from numeric ability. The former can be captured better by considering the frequency of nonintuitive responses. As such, we calculated an Inhibitory Control Score using the number of intuitive responses given rather than the number of correct responses to have a better measurement of the inhibitory control response (from here onward, "CRT-ICS").

Because the decision-task proposed to participants consists of a high numerical component, and because of the typical association found between the CRT and numerical skills, in our study we also used the Numeracy test as a potential confounding variable.

Numeracy is a cognitive skill recently and widely studied for its characteristics, and for its predictive power in different fields. It has been defined as a quantitative literacy or mathematical literacy and there are several definitions in literature related to aspects of this construct. As a test of mathematical and probabilistic reasoning we used the 8-item numeracy scale developed by Weller et al. ([59]). This scale consists of eight items, two of which are taken from the CRT scale by Frederick ([2]).

9. Method

9.1. Participants

The participants were 54 university students (34 female, 20 males; mean age= 23.6 years, SD=4.5). The students received 3 euros for participation and from a minimum of 3 euros to a maximum of 11.5 euros on the basis of their CRTs' performance. The participants who completed this experiment also completed the experiment reported by Dorigoni, Rajsic, & Bonini [60].

9.2. Procedure

Each participant, after reading the initial instructions and filling the informed consent, first participated in the decision-making section, and afterwards took four different CRTs and one numeracy test without any time limit. We recorded participants' eye movements only during the decision-making part of the experiment.

Decision-making task: choosing where to purchase.

Based on the results of the Pilot Study, three commercial packages were used in the decision-making task: the “washing machine”, "holiday by the sea" and “gym membership” packages. The following table (Table 2) summarizes the four price dimensions used for each package.

Attribute	Holiday by the sea	Washing machine	Gym membership
Similar	Room charge	Washing machine cost	Gym membership cost
Similar	Tourist tax	Transport cost	Registration card cost
Similar	Breakfast	Mounting cost	Discount
Dissimilar	Credit card fee	Extended warrant	Sauna

Table 2: Four price dimensions used for each package.

In each decision scenario, participants were told that the same package was available in two different online shops, "X website" and "Y website", and they had to decide which website they would purchase from by clicking on the appropriate button. The two websites were offering the same package but with a different distribution of the four values on the same four price-dimensions. The price-dimensions were displayed in a 2 by 4 table, with the rows corresponding to the two websites and the columns to the four price dimensions. In Figure 4, an example of decision scenario is shown, specifically the printer package that was used only as a training trial for the participants.

In sum, for each decision scenario, there were 4 price dimensions for each alternative package, which we will refer to generally as s1, s2, s3, and d which was critical to measuring attention to semantic pricing as will be explained shortly.

Printer				
	Printer's price	Transport	Installation	Extended warranty
X website	580 €	15 €	5%	55 €
Y website	460 €	30 €	30 %	52 €

Figure 4: Decision-making part: printer package scenario that was used as a training trial for the participants.

9.3. Experimental Design

Position of the "d" price in the decision table. The dissimilar price "d" was displayed either in the fourth (condition 1), the first (condition 2) or the third column (condition 3). Participants were randomly allocated to one of these three conditions. We manipulated this between-subject factor, to control whether the hypothesized "semantic-neglect" would be affected by the relative position of price "d" in the decision table (e.g., the price "d" might be easily neglected when displayed in the last than in the first column).

Numerical values for the "d" price in the decision table. To control whether the hypothesized "semantic-neglect" would be affected by the type of value used to express the dissimilar price "d", we built three versions for each package (holiday, washing machine and gym). In one version, only the price "d" was expressed as a percentage, in the second version only the price "s3" was expressed as a percentage, and in the third version both prices "d" and "s3" were expressed as percentages. Thus, participants were presented with nine decision scenarios: three versions for each of the three packages.

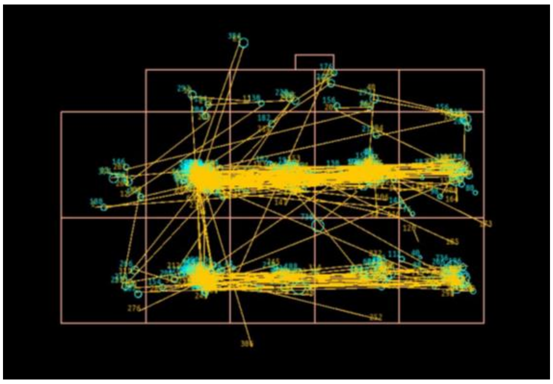
Criticality of the "d" price. For each of the nine decision scenarios, one website had a better promotion than the other (i.e., the full net price of the package was lower). It never happened that the two websites were offering the same package at the same net price. Specifically, in 4/9 of the trials the "X website" had the better promotion, and in the remaining trials the "Y website" was the better option. Furthermore, in 4/9 decision scenarios the values of the "d" prices were essential to choosing the cheapest option, as the values of the d dimension reversed the option rankings when included in the total price compared to when their values were excluded ("critical trials"). In the 5/9 trials the same option was cheapest regardless of whether the "d" price was included or not in the total price ("uncritical trials"). This manipulation is necessary to test whether the neglect of price "d" affects the quality of choice. If it does it, then the neglect of the "d" price in the critical trials would lower decision quality.

In sum, each participant was presented with 9 decision scenarios. She saw 3 different package scenarios, and for each package scenario she evaluated 3 different versions depending on which price dimension was expressed in percentage ("d", "s3" or both). Last, depending on the between-subject condition ("position of the price "d"), in all 9 decision scenarios, she saw the "d" price in the same column (first, third or fourth). The nine decision scenarios were randomly presented to the participant.

Eye-tracking

Eye movement data were recorded using an Eye Link 1000 Plus Binocular Tower Mount, which provides a data acquisition at up to 2000 Hz. The participants completed the decision-making task in the Psychtoolbox interface on a Dell computer with a 23 inches screen.

We defined fifteen areas of interest (AOIs), as shown in Figure 5 (in orange). In the analysis we consider the eight areas that contain the price values. These areas include all relevant information needed by the participants to make their decisions based on a rational integration of prices. Non-numerical information (e.g., the type of product) was not considered in the analysis because it was made clear at the beginning of the experiment that the two shops were delivering the same identical product. In the Figure 5 it is also shown the output of the eye-tracking measures



of a participant in a trial; the yellow lines represent the movements between different locations, while the numbers in light blue represent the duration of each fixation [61].

Figure 5: Fixations and saccades of a participant in a trial.

Tests of Cognitive Reflection and Numeracy

We used four different Cognitive Reflection Tests (CRT) in order to have a stronger measure of this construct, and to address the problem of the diffusion of the answers to the original CRT by Frederick ([2]). The original CRT by Frederick was a three-item test. Given its increasing popularity (it is becoming common knowledge among the university student population), many other tests were developed as alternatives.

Participants completed the original Frederick's test, and the tests developed by Toplak et al., ([36]), Primi et al. ([62]) and Thomson and Oppenheimer ([57]).

We analyze three CRTs and the original one by Frederick; some of these cognitive reflection tests added the items to the original test. Toplak and colleagues added four items (CRT4) to the original three-item test (CRT3) with the result being a seven-item test (CRT7). Primi and colleagues investigated not only the psychometric properties of the original CRT but also a new version of it with three new items (CRT-L). Thomson and Oppenheimer developed a four-item test (CRT-2) in order to increase the pool of available questions. The final cognitive reflection score used in the analysis is the proportion of the nonintuitive answers divided by the total number of answers. The three Frederick's questions were considered only once. Thus, each participant answered in total to 14 CRT questions.

As a test of mathematical and probabilistic reasoning, we used the 8-item numeracy scale developed by Weller et al. ([59]). This scale consisted of eight items, two of which were taken from the CRT scale by Frederick.

10. Hypotheses

The first hypothesis is related to the choice accuracy.

HYPO 1 (Choices): We expected that choice accuracy will be lower with low CRT-ICS than high CRT-ICS people. The lower choice accuracy of low CRT-ICS people could be due to two alternative visual attention processes that are described below:

HYPO 2.1: "Neglect of the semantically unrelated price-dimension 'd'". Low CRT-ICS people barely look at the "price-d" compared to high CRT-ICS people because of the semantic grouping of price-dimensions. Thus, we expected that the average dwell time spent on the dissimilar price "d" is nearly to zero for low CRT-ICS people.

Because of the semantic grouping of price-dimensions, we should also find a lower choice accuracy of the low CRT-ICS people compared to High CRT-ICS people, but only in the critical trials, that is when the integration of the price “d” is crucial to detect the best package deal. Said differently, we expected to find an interaction between the factor “CRT-ICS” and the factor “Type of trial” (critical vs. uncritical).

HYPO 2.2: “Superficial processing of all displayed price dimensions”. From this hypothesis, both high and low CRTICS people should look at all price-dimensions (e.g., no neglect of the dissimilar price “d”), but low CRTICS people should be faster than high CRTICS people. This would be due to the lower inhibition component, and heuristic price-information processing that characterize impulsive people. This result should be found regardless of type of price-dimension (similar VS dissimilar).

The following analyses were planned to test the above hypotheses. The first set of analysis have the choice accuracy (1=accurate; 0=inaccurate) as dependent variable:

- An analysis of the choice accuracy to control for the effect of the position of the price ‘d’ (Table 3 first model), and its expression type (e.g, ‘d’ as a percentage or not) on choice accuracy (Table 3 second model);
- An analysis where the dependent variable was choice accuracy and the predictors were CRT-ICS score, Type of trial (critical vs. uncritical), numeracy, and the interaction between CRT-ICS and Type of trial (Table 4).

The second set of analysis assessed differences in visual attention using a linear mixed model with the dependent variable the average dwell time across the eight prices, predicted by CRT-ICS scores, numeracy, and "Type of price dimension” (similar VS dissimilar).

11. Results

In order to use the purest measure of numeracy, we removed the two Frederick’s questions from the Numeracy test, obtaining the 6-item Numeracy, because it is less related to cognitive reflection (Table A4 in the Appendix A). We also decided to analyze the CRTICS instead of the classic CRT for the reasons mentioned earlier and because it is less related to the numeracy, although they are still significantly correlated.

Due to these points, the following analysis will consider the CRT-ICS and the 6-item Numeracy.

Choice

To control for the effect of the “position of price ‘d”’ (whether it is displayed in the fourth, in the first or in the third column), and “its expression type” (whether ‘d’ is expressed by a percentage or an absolute value) on choice accuracy, we computed two glm models with the binary dependent variable "choice accuracy" that indicates whether a participant chooses ("1") or not ("0") the cheapest package. The independent variables are the “position of price ‘d”’ (referent category, “last column"), and the “type of expression of ‘d”’ (referent category, “ ‘d’ in percentage").

Table 3: Choice accuracy as a function of the position and the percentage

	Dependent variable:	
	Choice accuracy	
	(1)	(2)
d in the first column	−0.343	
	(0.255)	
d in the third column	−0.369	
	(0.266)	
d and s in perc		−0.212
		(0.238)
s in perc		−0.053
		(0.244)
Constant	0.750***	0.610***
	(0.204)	(0.199)
Observations	486	486
Log Likelihood	−319.414	−320.219
Akaike Inf. Crit.	648.829	650.438
Note:	*p<0.1; **p<0.05; ***p<0.01	

As shown in Table 3, choice accuracy is not affected neither by the position of price “d” nor by its numerical expression. Thus, for the further analysis, we aggregated the data because these two factors have no effect on choice accuracy.

HYPO 1: Choice accuracy and CRT

We studied the relationship between choice accuracy and CRT-ICS by means of a glm model with the binary dependent variable "choice accuracy" that indicates whether a participant chooses ("1") or not ("0") the cheaper package.

Here, we test the effect of the CRT-ICS on the choice accuracy while controlling for the numeracy level. If we had not controlled for numeracy, then part of the effect of CRT-ICS could actually be the result of high CRT people with high numeracy, and those who have higher numeracy tend to be accurate in calculation. We also wanted to see if there is an interaction effect between CRT-ICS and the factor “type of trial” (critical/uncritical) through a dummy variable (referent category: “uncritical”), in order to understand if the effect of CRT-ICS is moderated by the type of trial.

We found that the CRT-ICS has a significant and positive effect on choice accuracy ($b_{M2,\ CRT-ICS}$ = 1.489, SE = 0.72, z = 2.051, p = 0.040) also when it is controlled for numeracy (Table 4, model 2). Thus, HYPO 1 is confirmed, as low CRT-ICS people make less accurate choices than high CRT-ICS people.

Table 4: Choice accuracy of the decision-making task

	Dependent variable:	
	Choice accuracy	
	(1)	(2)
6-item Numeracy	1.271***	0.524
	(0.436)	(0.586)
CRT-ICS		1.490**
		(0.726)
Critical		0.111
		(0.631)
CRT-ICS:Critical		−0.818
		(0.827)
Constant	−0.297	−0.432
	(0.376)	(0.467)
Observations	486	486
Log Likelihood	−305.341	−303.026
Akaike Inf. Crit.	618.682	620.051
Bayesian Inf. Crit.	635.427	649.355
Note:	*p<0.1; **p<0.05; ***p<0.01	

The interaction of “CRT-ICS” with “Type of trial” is not significant. That is, the higher the CRT-ICS the higher the choice accuracy, regardless of type of trial. Said differently, high CRT-ICS people are more accurate than low CRT-ICS people both when the integration of the price “d” with the other prices is necessary to detect the cheaper package, and when it is not. This finding does not confirm HYPO 2.1 (semantic-neglect hypothesis).

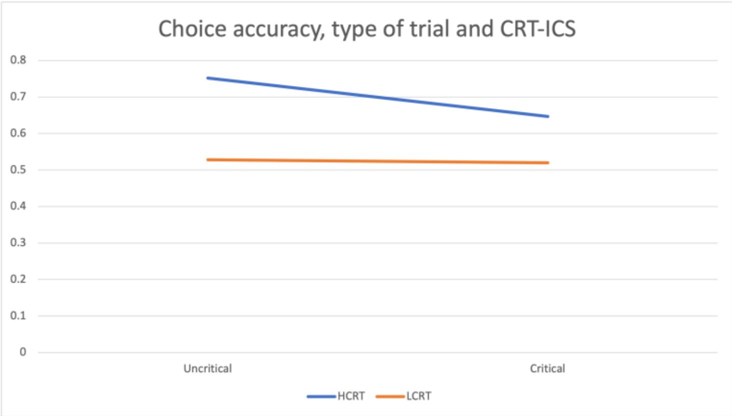


Figure 6: Choice accuracy as a function of type of trial (uncritical VS critical) and CRT-ICS.

In the Figure 6 we can see the average values of choice accuracy of the two groups (low and high CRT-ICS considering the median) in the uncritical and critical trials. People with high CRT choose the best deal 75% of the time in the uncritical trials (when the dissimilar attribute is not critical to the choice) and 65% of the time in the critical trials. People with low CRT-ICS choose the best deal 53% of the time in the uncritical trials and 52% of the time in the critical trials. The fact that low CRT-ICS people have the same level of choice accuracy between the two types of trial does not confirm the semantic neglect hypothesis.

Visual Attention

To understand why high CRT-ICS people choose systematically better than low CRT-ICS people, we analyze the gaze behavior of these two groups to see whether fixation times differ as a function of CRT, type of price dimension (similar vs. dissimilar), and their interaction. Specifically, we computed the average dwell time for each AOI as a measure of the visual attention allocated by the participant to each of the eight displayed prices [63].

To test the “semantic-neglect” hypothesis (HYPO 2.1: CRT-ICS moderates the neglect of price "d"), we need to understand if the effect of CRT-ICS on average dwell time is different between similar and dissimilar prices. This hypothesis entails an interaction between CRT-ICS and type of price dimension because the difference in dwelling times between the two groups should be semantically bounded. Said differently, only low CRT-ICS people should neglect the "d" price.

However, from the "Superficial processing of all displayed price dimensions" (HYPO 2.2), due to the lower inhibition component and the heuristic price-information processing, we expected that high CRT-ICS people have longer dwell times than low CRT-ICS people, regardless of the type of price information (similar Vs. dissimilar price dimensions). Said differently, we expected to find a main effect of CRT-ICS but without its interaction with the type of price dimension. This means that high CRT people simply look longer and more frequently than low CRT-ICS people at each price dimension, regardless of its semantic relatedness to the purchase.

HYPO 2.1 and 2.2: Average dwell time and individual differences

In table 5, the average dwell time across all eight displayed prices (for each trial and each participant) is reported. The model 1 in Table 5 (first column) tests the effect of "CRT-ICS" on average dwell time controlling for "Numeracy". If we had not controlled for numeracy, then part of the effect of CRT-ICS on visual attention could be the result of the influence of numeracy. Results show that both factors are statistically significant.

Keeping constant the Numeracy level, participants spent on average almost 7 seconds more in looking at prices for every increasing level in the CRT-ICS. The CRT-ICS factor is positively statistically significant ($b_{M1,\ CRT-ICS}$ = 6634.2, SE = 624.4, t = 10.625, $p < 0.001$).

Keeping constant the CRT-ICS level, participants spent on average almost 1.2 seconds less in looking at prices for every increasing level in "Numeracy". The Numeracy factor is negatively statistically significant ($b_{M1,\ \mathrm{NUM}}$ = -1203.6, SE = 604.7, t = -1.990, p = 0.046). This could be because higher numerate people are quicker in making calculations.

In the second model, in order to test the HYPO 2.1 and 2.2 we add the interaction between CRT-ICS and the type of price-dimension (similar or dissimilar), because we want to test if the effect of CRT-ICS is moderated by the type of price-dimension. Results show that, again, the main factor CRT-ICS is positively statistically significant while controlling for numeracy level ($b_{M2,\ CRT-ICS}$ = 7199.1, SE = 671.9, t = 10.715, $p < 0.001$).

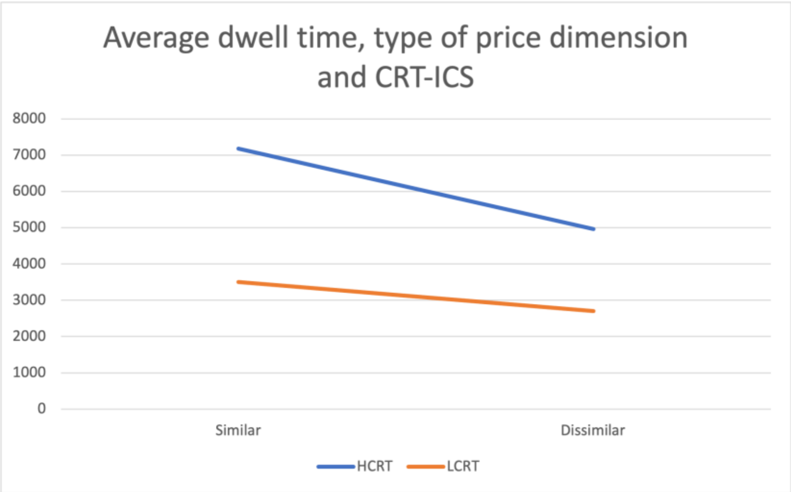
Table 5: Dwell time on AOIs

	Dependent variable:	
	Dwell time on AOIs	
	(1)	(2)
6-item Numeracy	$-1{,}203.561^{**}$	$-1{,}203.561^{**}$
	(604.739)	(601.465)
CRT-ICS	$6{,}634.158^{***}$	$7{,}199.135^{***}$
	(624.388)	(671.891)
Dissimilar		-363.484
		(596.310)
CRT-ICS:Dissimilar		$-2{,}259.906^{**}$
		(1,025.965)
Constant	$2{,}393.490^{***}$	$2{,}484.361^{***}$
	(314.835)	(346.806)
Observations	3,888	3,888
R^2	0.042	0.053
Adjusted R^2	0.042	0.052
Residual Std. Error	6,725.325 (df = 3885)	6,688.910 (df = 3883)
F Statistic	85.294^{***} (df = 2; 3885)	54.217^{***} (df = 4; 3883)

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

However, the effect of CRT-ICS is moderated by the type of price-dimension ($b_{M2,\ CRT-ICS*TYPE}$ = -2259.9, SE = 1026.0, t = -2.203, p = 0.027). This means that high CRT-ICS participants spend more time than low CRT-ICS participants at looking at both similar and dissimilar prices, but that difference is more pronounced for the similar prices.

In the Figure 7 we can see the average dwell time of the two groups (low and high CRT-ICS) spent on the similar VS dissimilar attributes. People with high CRT-ICS spent on average around 7.2 seconds on the similar attributes and 5 seconds on the dissimilar attributes while people with low CRT spent 3.5 seconds on the similar attributes and 2.7 seconds on the dissimilar attributes. Considering the regression model mentioned above, people with higher CRT-ICS spend more time on both, similar and dissimilar attributes, but the gap between low and high CRT in terms of time spent on AOIs is higher in the similar prices compared to the dissimilar prices. The fact that fixation times of low CRT-ICS are around 3 seconds for both types of price-dimension leads us to reject HYPO 2.1 (neglect of dissimilar price dimension). The fact that fixation times of low CRT-ICS are systematically lower than high CRT-ICS confirms HYPO 2.2 (superficial processing of all displayed price dimensions) [1].



1 We also computed a mediation analysis (even if the sample size is too small for a mediation analysis N = 54) that shows the effect of the CRT on the choice accuracy, and it is mediated neither by dwell nor by numeracy. Thus, the process underlying the correct choice (which is significantly influenced by the CRT regardless of the level of numeracy and dwell) would appear to be caused by other intervening variables.

Figure 7: Average dwell time as a function of type of price dimension (similar VS dissimilar) and CRT-ICS.

12. Discussion

Reported results show that cognitive reflection affects the quality of choice. An impulsive consumer has nearly a 20% loss in choice accuracy compared to a reflective consumer (approximately 50% vs. 70%, respectively). Said differently, an impulsive consumer selects the best commercial offer nearly at a chance rate. This finding holds also when consumer's numeracy ability is statistically controlled. This suggests that cognitive reflection is a distinct driver of choice compared to numeracy, even if the measures of these two constructs are correlated. Similar findings are reported in Graffeo et al., [19] where choice accuracy was lower for low than high CRT people in the context of a consumer choice when the numeracy ability was statistically controlled.

The poorer decision quality of impulsive consumers seems to be due to their distinctive way of looking at the price-information. They look at all displayed prices but very quickly: The average dwelling time for an impulsive consumer is around 3 seconds compared to the 6 seconds of the more reflective one. This finding suggests that impulsive consumers tend to form a quick impression of what seems to be the best offer. Again, quite similar findings are reported in Graffeo et al., [19] where low CRT people had both lower fixations times and fixations counts at the initial undiscounted displayed prices compared to high CRT people. This shallow visual search pattern not surprisingly holds a poorer decision quality; the attention allocation of low cognitive participants is not explained by the semantic neglect bias hypothesis but by the superficial price-information processing hypothesis. Lazy, not biased as the authors Pennycook and Rand [67] explained the lack of reasoning associated to a superficial judgment. A superficial visual exploration coherent with the results found by Del Missier and colleagues [68] where the monitoring/inhibition dimension of executive functions was significantly related to CRT; cognitive reflection defined as the ability to "favor analytic over intuitive processes when needed in judgment and decision making," ([69] pp. 274–275).

Theoretically, these findings add to previous evidence to clarify the distinct role of impulsiveness both at the information-processing level and the quality of choice as well as at the relationship between these two aspects of decision-making. Reported findings are also useful for informing current issues in marketing and business law. Private marketing often uses so called "drip pricing" techniques in which only a part of a product or service's total price is first advertised, with the total price amount provided only at the end of the buying process. For example, in the context of online ticketing purchase, the consumer is first provided with the cost of the ticket (e.g., 10€) and only later with the administration fee (5€), and eventually the fee for the credit card payment (5€). This pricing technique is very effective in affecting consumer choices as reported in a set of empirical studies in 2010 by the UK Office of Fair Trading (OFT); they concluded that drip pricing had the greatest potential to mislead consumers, and "complex [price] offers" were ranked third ([70]).

Furthermore, from the European Directive on Unfair Commercial Practices, UCPD (European Parliament and Council, 2005), the "drip-pricing" technique is considered a misleading business practice (see, for example, the Italian Authority of Competition and Market (AGCM) legal cases on "MY AIR" (Case PS 168), and on "TRANSAVIA" (Case PS 9670)). The reason is that the consumer is induced to accept an economic transaction because under these circumstances (e.g., with a drip-pricing format where the initial price is very low) is induced in error (e.g., induced to expect that this transaction is offering the best price in the market. See European Commission (2016) for a guideline document related to the application of the UCPD where consumer expectation is also considered).

Reported findings cast doubts on whether this level of consumer protection should be extended to the "multi-dimensional pricing" technique where all prices are displayed at once (and not temporarily distributed as with the "drip-pricing format"). Our study, in fact, found that impulsive consumers are associated with poorer performance in those contexts (see [71] for a review on how partitioned pricing works, the mechanisms by which it exerts its impact, and the appropriate areas where the practice may need regulation to protect consumers).

Last, the finding that the same pricing technique is detrimental mostly for a specific consumer category (e.g., the impulsive consumer) contributes to the open debate in the business law on what is meant with "average consumer", and which level of protection she ultimately deserves.

The Italian "Consiglio di Stato" has recently approved an order (CONSIGLIO DI STATO; sezione VI; ordinanza 10 ottobre 2022, n. 8650) to address the Court of Justice of the European Union (CGUE) in order, amongst other aspects, to determine whether the current definition of "average consumer" is adequate in the light of recent findings in the behavioral economics and cognitive psychology literatures ([72] for a comment). Our findings empirically corroborate that legal argument because under the same "pricing-scheme" there were huge differences in the quality of choices as a function of consumer impulsiveness. So, it seems unplausible to refer to the notion of an abstract "average consumer" under these circumstances.

Appendix A. Appendix

Typology		room charge	tourist tax	breakfast	credit card fee	discount	wi-fi fee
similar attribute	room charge	1	0.789	0.895	0.053	0.368	0.316
similar attribute	tourist tax	0.789	1	0.789	0.211	0.368	0.316
similar attribute	breakfast	0.895	0.789	1	0.105	0.421	0.421
dissimilar attribute	credit card fee	0.053	0.211	0.105	1	0.579	0.316
rejected attribute	discount	0.368	0.368	0.421	0.579	1	0.263
rejected attribute	wi-fi fee	0.316	0.316	0.421	0.316	0.263	1

Table A1. Pilot study result for the holiday task.

Typology		washing machine cost	transport cost	mounting cost	extended warranty	disposal cost	appropriate detergent
similar attribute	washing machine cost	1	0.474	0.579	0.105	0.316	0.632
similar attribute	transport cost	0.474	1	0.684	0.421	0.316	0.368
similar attribute	mounting cost	0.579	0.684	1	0.368	0.263	0.474
dissimilar attribute	extended warranty	0.105	0.421	0.368	1	0.474	0.316
rejected attribute	disposal cost	0.316	0.316	0.263	0.474	1	0.211
rejected attribute	appropriate detergent	0.632	0.368	0.474	0.316	0.211	1

Table A2. Pilot study result for the washing machine task.

Typology		gym membership cost	registration card cost	discount	sauna	locker cost	doctor's appointment
similar attribute	gym membership cost	1	0.842	0.632	0	0.105	0.211
similar attribute	registration card cost	0.842	1.000	0.579	0.053	0.211	0.263
similar attribute	discount	0.632	0.579	1.000	0.158	0.211	0.158
dissimilar attribute	sauna	0	0.053	0.158	1.000	0.526	0.579
rejected attribute	locker cost	0.1052632	0.211	0.211	0.526	1.000	0.263
rejected attribute	doctor's appointment	0.2105263	0.263	0.158	0.579	0.263	1

Table A3. Pilot study result for the gym membership task.

	8 item numeracy	6 item numeracy	CRT -ICS	CRT
8 item numeracy				
6 item numeracy	0.94***			
CRT-ICS	0.77***	0.64***		
CRT	0.82***	0.70***	0.9***	
Mean	0.65	0.68	0.53	0.71
SD	0.26	0.25	0.24	0.21

Table A4. Intercorrelation among Numeracy and CRT.

References

[1] K. E. Stanovich, R. F. West, Individual differences in reasoning: Implications for the rationality debate?, Behavioral and brain sciences 23 (5) (2000) 645–665.

[2] S. Frederick, Cognitive reflection and decision making, Journal of Economic perspectives 19 (4) (2005) 25–42.

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

[4] E. I. Hoppe, D. J. Kusterer, Behavioral biases and cognitive reflection, Economics Letters 110 (2) (2011) 97–100.

[5] J. M. Liberali, V. F. Reyna, S. Furlan, L. M. Stein, S. T. Pardo, Individual differences in numeracy and cognitive reflection, with implications for biases and fallacies in probability judgment, Journal of behavioral decision making 25 (4) (2012) 361–381.

[6] D. Kahneman, Thinking, fast and slow, Farrar, Straus and Giroux: New York., 2011.

[7] C. Alós-Ferrer, M. Garagnani, S. Hügelschäfer, Cognitive reflection, decision biases, and response times, Frontiers in psychology 7 (2016) 1402.

[8] C. Alós-Ferrer, S. Hügelschäfer, Faith in intuition and cognitive reflection, Journal of Behavioral and Experimental Economics 64 (2016) 61–70.

[9] B. Moritz, E. Siemsen, M. Kremer, Judgmental forecasting: Cognitive reflection and decision speed, Production and Operations Management 23 (7) (2014) 1146–1160.

[10] J. Zonca, G. Coricelli, L. Polonio, Gaze patterns disclose the link between cognitive reflection and sophistication in strategic interaction, Judgment and Decision making 15 (2) (2020) 230–245.

[11] I. Otero, J. F. Salgado, S. Moscoso, Cognitive reflection, cognitive intelligence, and cognitive abilities: A meta-analysis, Intelligence 90 (2022) 101614.

[12] J. M. Liberali, V. F. Reyna, S. Furlan, L. M. Stein, S. T. Pardo, Individual differences in numeracy and cognitive reflection, with implications for biases and fallacies in probability judgment, Journal of behavioral decision making 25 (4) (2012) 361–381.

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

[14] O. H. Sari, Theory of planned behaviour in marketing: Cognitive consideration on purchase decision, Golden Ratio of Mapping Idea and Literature Format 2 (1) (2022) 01–07.

[15] G. Campitelli, M. Labollita, Correlations of cognitive reflection with judgments and choices, Judgment and Decision making 5 (3) (2010) 182–191.

[16] M. Sirota, M. Juanchich, Role of numeracy and cognitive reflection in Bayesian reasoning with natural frequencies., Studia Psychologica 53 (2) (2011).

[17] E. Peters, I. P. Levin, Dissecting the risky-choice framing effect: Numeracy as an individual-difference factor in weighting risky and riskless options, Judgment and Decision making 3 (6) (2008) 435–448.

[18] A. Sobkow, A. Olszewska, J. Traczyk, Multiple numeric competencies predict decision outcomes beyond fluid intelligence and cognitive reflection, Intelligence 80 (2020) 101452.

[19] M. Graffeo, L. Polonio, N. Bonini, Individual differences in competent consumer choice: the role of cognitive reflection and numeracy skills, Frontiers in psychology 6 (2015) 844.

[20] A. Dorigoni, An eye tracking exploration of cognitive reflection in consumer decision-making (2019).

[21] A. Dorigoni, L. Polonio, M. Graffeo, N. Bonini, Analisi di movimenti oculari nella risoluzione di problemi commerciali: Il ruolo delle abilità numeriche e dell'impulsività cognitiva, FrancoAngeli, 2022.

[22] T. C. Burnham, D. Cesarini, M. Johannesson, P. Lichtenstein, B. Wallace, Higher cognitive ability is associated with lower entries in a p-beauty contest, Journal of Economic Behavior & Organization 72 (1) (2009) 171–175.

[23] O. Bergman, T. Ellingsen, M. Johannesson, C. Svensson, Anchoring and cognitive ability, Economics Letters 107 (1) (2010) 66–68.

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

[25] H. Estelami, The effect of price presentation tactics on consumer evaluation effort of multi-dimensional prices, Journal of Marketing Theory and Practice 11 (2) (2003) 1–16.

[26] H. Estelami, Strategic implications of a multi-dimensional pricing environment, Journal of Product & Brand Management 12 (5) (2003) 322–334.

[27] D. A. Hensher, The relationship between bus contract costs, user perceived service quality and performance assessment, International Journal of Sustainable Transportation 8 (1) (2014) 5–27.

[28] V. G. Morwitz, E. A. Greenleaf, E. J. Johnson, Divide and prosper: consumers' reactions to partitioned prices, Journal of marketing research 35 (4) (1998) 453–463.

[29] A. T. Abraham, R. W. Hamilton, When Does Partitioned Pricing Lead to More Favorable Consumer Preferences?: Meta-Analytic Evidence, Journal of Marketing Research 55 (5) (2018) 686–703.

[30] E. Rosch, C. B. Mervis, Family resemblances: Studies in the internal structure of categories, Cognitive psychology 7 (4) (1975) 573–605.

[31] E. Rosch, Principles of Categorization Erlbaum, Hillsdale, NJ (1978).

[32] L. W. Barsalou, Context-independent and context-dependent information in concepts, Memory & cognition 10 (1) (1982) 82–93.

[33] P. W. Henderson, R. A. Peterson, Mental accounting and categorization, Organizational Behavior and Human Decision Processes 51 (1) (1992) 92–117.

[34] E. R. Evers, A. Imas, C. Kang, On the role of similarity in mental accounting and hedonic editing., Psychological review 129 (4) (2022) 777.

[35] R. H. Thaler, Mental accounting matters, Journal of Behavioral decision making 12 (3) (1999) 183–206.

[36] M. E. Toplak, R. F. West, K. E. Stanovich, Assessing miserly information processing: An expansion of the Cognitive Reflection Test, Thinking & Reasoning 20 (2) (2014) 147–168.

[37] E. T. Cokely, A. Feltz, S. Ghazal, J. N. Allan, D. Petrova, R. Garcia-Retamero, Decision making skill: From intelligence to numeracy and expertise, Cambridge handbook of expertise and expert performance

(2018) 476–505.

[38] B. M. Velichkovsky, From levels of processing to stratification of cognition, Stratification in cognition and consciousness (1999) 203–235.

[39] B. M. Velichkovsky, Heterarchy of cognition: The depths and the highs of a framework for memory research, Memory 10 (5-6) (2002) 405–419.

[40] B. M. Velichkovsky, A. Rothert, M. Kopf, S. M. Dornhöfer, M. Joos, Towards an express-diagnostics for level of processing and hazard perception, Transportation Research Part F: Traffic Psychology and Behaviour 5 (2) (2002) 145–156.

[41] S. Eivazi, R. Bednarik, Predicting problem-solving behavior and performance levels from visual attention data, in: Proc. workshop on eye gaze in intelligent human machine interaction at IUI, 2011, pp. 9–16.

[42] A. Glöckner, A.-K. Herbold, 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) (2011) 71–98.

[43] J. L. Orquin, M. P. Bagger, S. M. Loose, Learning affects top down and bottom up modulation of eye movements in decision making, Judgment and Decision making 8 (6) (2013) 700–716.

[44] S. Fiedler, A. Glöckner, The dynamics of decision making in risky choice: An eye-tracking analysis, Frontiers in psychology 3 (2012) 335.

[45] H. M. Kim, L. Kachersky, Dimensions of price salience: a conceptual framework for perceptions of multi-dimensional prices, Journal of Product & Brand Management 15 (2) (2006) 139–147.

[46] E. F. Cataldo, R. M. Johnson, L. A. Kellstedt, L. W. Milbrath, Card sorting as a technique for survey interviewing, Public Opinion Quarterly 34 (2) (1970) 202–215.

[47] N. Nurmuliani, D. Zowghi, S. P. Williams, Using card sorting technique to classify requirements change, in: Proceedings. 12th IEEE International Requirements Engineering Conference, 2004., 2004, pp. 240–248.

[48] T. S. Tullis, Designing a menu-based interface to an operating system, in: ACM SIGCHI Bulletin, Vol. 16, 1985, pp. 79–84.

[49] K. Frederickson-Mele, Usability testing an intranet prototype shell, in: Web Usability Workshop at Chi, Vol. 97, 1997.

[50] T. S. Tullis, Using card-sorting techniques to organize your intranet, Intranet Journal of Strategy & Management 1 (2003) 1–9.

[51] T. Tullis, L. Wood, How many users are enough for a card-sorting study Proceedings of the Usability Professionals Association Conference, Minneapolis, MN (2004).

[52] T. Tullis, L. Wood, How many users are enough for a card-sorting study, in: Proceedings UPA, Vol. 2004, 2004.

[53] A. Tversky, D. Kahneman, The framing of decisions and the psychology of choice, science 211 (4481) (1981) 453–458.

[54] K. Deibel, R. Anderson, R. Anderson, Using edit distance to analyze card sorts, Expert Systems 22 (3) (2005) 129–138.

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

[56] A. Sinayev, E. Peters, Cognitive reflection vs. calculation in decision making., Frontiers in Psychology 6 (2015) 532–532.

[57] K. S. Thomson, D. M. Oppenheimer, Investigating an alternate form of the cognitive reflection test, Judgment and Decision Making 11 (1) (2016) 99.

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

[59] J. A. Weller, N. F. Dieckmann, M. Tusler, C. K. Mertz, W. J. Burns, E. Peters, Development and testing of an abbreviated numeracy scale: A Rasch analysis approach, Journal of Behavioral Decision Making 26 (2) (2013) 198–212.

[60] A. Dorigoni, J. Rajsic, N. Bonini, Does cognitive reflection predict attentional control in visual tasks?, Acta Psychologica 226 (2022) 103562.

[61] S. T. Fiske, M. A. Pavelchak, Category-based versus piecemeal-based affective responses: Developments in schema-triggered affect. (1986).

[62] C. Primi, K. Morsanyi, F. Chiesi, M. A. Donati, J. Hamilton, The development and testing of a new version of the cognitive reflection test applying item response theory, Journal of Behavioral Decision Making (2015).

[63] Z. Sharafi, T. Shaffer, B. Sharif, Y.-G. Guéhéneuc, Eye-tracking metrics in software engineering, in: 2015 Asia-Pacific Software Engineering Conference (APSEC), 2015, pp. 96–103.

[64] P. Sleboda, J. Sokolowska, Measurements of rationality: Individual differences in information processing, the transitivity of preferences and decision strategies, Frontiers in psychology 8 (2017) 1844.

[65] J. Stuyck, European consumer law after the Treaty of Amsterdam: Consumer policy in or beyond the internal market?, Common Market Law Review 37 (2) (2000).

[66] W. H. V. Boom, A. Garde, O. Akseli, Introduction to'The European Unfair Commercial Practices Directive', Van Boom, Garde & Akseli, The European Unfair Commercial Practices DirectiveImpact, Enforcement Strategies and National Legal Systems (2014) 1–18.

[67] G. Pennycook and D. G. Rand. "Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning." *Cognition* 188 (2019): 39-50.

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

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

[70] The Office of Fair Trading (OFT) (2010) Information Policy Team, The National Archives, Kew, London https://www.gov.uk/government/organisations/office-of-fair-trading

[71] Greenleaf, E. A., Johnson, E. J., Morwitz, V. G., & Shalev, E. (2016). The price does not include additional taxes, fees, and surcharges: A review of research on partitioned pricing. *Journal of Consumer Psychology*, *26*(1), 105-124.

[72] C. Bona, Carlo, N. Bonini, Tutela del consumatore e nuovi paradigmi scientifici: scienze cognitive e neuroscienze varcano la soglia di palazzo Spada, Note to: Consiglio di Stato, Sezione VI, ordinanza 10 ottobre 2022 n. 8650 (rinvio pregiudiziale alla Corte di Giustizia dell'Unione Europea), in IL FORO ITALIANO, Parte 03, Colonna 542, Fascicolo 11 (2022).