Original Article

Cognitive Reflection Effects on Time Discounting

Michał Białek[1,2] and Przemysław Sawicki[1]

[1]Centre for Economic Psychology and Decision Sciences, Koźminski University, Warsaw, Poland
[2]Department of Psychology, University of Waterloo, Ontario, Canada

Abstract: In this work, we investigated individual differences in cognitive reflection effects on delay discounting – a preference for smaller sooner over larger later payoff. People are claimed to prefer more these alternatives they considered first – so-called reference point – over the alternatives they considered later. Cognitive reflection affects the way individuals process information, with less reflective individuals relying predominantly on the first information they consider, thus, being more susceptible to reference points as compared to more reflective individuals. In Experiment 1, we confirmed that individuals who scored high on the Cognitive Reflection Test discount less strongly than less reflective individuals, but we also show that such individuals are less susceptible to imposed reference points. Experiment 2 replicated these findings additionally providing evidence that cognitive reflection predicts discounting strength and (in)dependency to reference points over and above individual difference in numeracy.

Keywords: cognitive reflection, query theory, construction of preferences, delay discounting, individual differences

[Author: please approve the integration of footnotes in the main text] With growing income and access to goods, the problem of impulsivity gains significance. Specifically, people are unable to delay gratifications and are overly willing to delay losses, which in turn lead to social problems such as debts, obesity, and addictions (MacKillop et al., 2011).

An expression of individual's impulsivity is the degree to which individuals can delay, labeled as discounting strength. Typical delay discounting task requires an individual to provide an equivalent which is subjectively indifferent to the evaluated payoff, for example, what is the immediate amount for which they would be indifferent to $5,000 to be received in 6 months. The smaller the equivalent, the greater the impulsivity. The equivalents serving as a basis for calculating the discounting strength can be experimentally obtained in several ways, for example, by directly asking participants about such equivalent (matching), or indirectly, based on choices between sooner and later payoffs (staircase and adjusting). For comparison of these methods see Hardisty, Thompson, Krantz, and Weber (2013). Typical findings suggest people are more impulsive when discounting gains over losses (Thaler & Shefrin, 1981), and for smaller over greater amounts (Benzion, Rapoport, & Yagil, 1989).

People often are inconsistent in their intertemporal choices. For example, they expect a greater bonus for receiving expected income later, compared to what they are willing to pay for receiving a similar income earlier (Loewenstein & Prelec, 1992). This finding, known as the direction effect, can be explained by query theory (Johnson, Häubl, & Keinan, 2007), which suggests that decision makers construct their preferences by making sequential internal queries about the available choice options. People initially address their queries to the option presented to them as the default, and these initial thoughts affect behavior more than do their subsequent thoughts. Indeed, Weber and colleagues (Appelt, Hardisty, & Weber, 2011; Weber et al., 2007) have shown that people have more impulsive thoughts in a delaying condition (when a smaller sooner payoff is presented as the default and is therefore considered first) than in an accelerating condition (when a larger later payoff is the default and is therefore considered first). Authors conclude that such order of thoughts explains the strength of discounting.

Cognitive science has researched individual differences in humans' sequential processing of information (De Neys & Bonnefon, 2013), suggesting that some people's thinking relies heavily on their first, intuitive thoughts, whereas others think more reflectively and consider more information. The willingness to override intuitive responses, termed cognitive reflection, is prominent in recent models of thinking (De Neys, 2014; Pennycook, Fugelsang, & Koehler, 2015).

https://doi.org/10.1027/1614-0001/a000254

There is some evidence that higher cognitive reflection is connected to weaker delay discounting (Frederick, 2005; Toplak, West, & Stanovich, 2011), suggesting a quantitative difference between less and more reflective individuals. Researchers do not explain this relationship beyond suggesting that cognitively reflective individuals cope better with impulsive thoughts, implicitly assuming that the cognitive process underlying delay discounting is the same in more and less reflective individuals and that the difference lies in the degree to which decision maker is affected by impulsive thoughts she considers.

However, less cognitive reflective people focus mostly on the first thoughts they consider, while more reflective people delay decisions until they have also considered other alternative(s). Hence, highly reflective people should be less susceptible to changes in problem formulation that affects the order of considered alternatives (acceleration vs. delaying). In other words, the difference between reflective and nonreflective individuals would be qualitative, affecting *how* they process information, and result in lesser variability of discount rates.

Understanding the effects of cognitive reflection of intertemporal choice in particular, and on decision-making in general, can be essential for improving quality of human lives. As mentioned above, individuals change their judgments and decisions according to the imposed reference point, or the perspective taken, and take contradictory actions when facing similar problems over time (Lee, Lee, Bertini, Zauberman, & Ariely, 2015). The ability to override or ignore irrelevant information like the reference point can support consistency in one's judgments and actions, decreasing susceptibility to manipulations such as framing of problems or nudges. Thus, highly reflective individuals can be said to be more rational, and their decisions may better reflect their true preferences (preferences they would display in optimal conditions, Baron, 2016).

Experiment 1

We predicted that cognitive reflection not only decreases impulsivity in delay discounting but also reduces the susceptibility to imposed reference points. We examined our hypothesis by examining variability in participants' discount rates in two tasks that differed according to which alternative was presented as the default. Specifically, in one task, participants were instructed that a payoff was to be received immediately (thus, a smaller sooner alternative was the default) and indicated how much bonus they would expect to delay this payoff (delaying condition). In the other task, participants were informed they would receive a payoff after a specific delay (i.e., a larger later payoff was the default) and indicated the greatest reduction of this payment that they would be willing to accept to receive the money immediately (accelerating condition). A greater difference between the obtained discount rates in accelerating and delaying gains would mean that preferences were affected by order of considered alternatives; conversely, no difference between the discount rates in these two tasks suggests that the preferences are relatively independent of the order of considered alternatives.

According to our reasoning, the impact of reference points on discount rates should be connected to cognitive reflection, so that more reflective individuals are less susceptible to changes in reference points.

Participants

Participants were native English speakers from the US, UK, and Canada (N = 213) recruited via Amazon Mechanical Turk and were compensated $0.80 for their effort. Besides the nationality, there was no other restriction for participating the experiment.

A subset of participants (n = 47; 22% of the sample) showed a non-monotonic decrease of value with a length of delay (e.g., a payoff delayed for 6 months was indicated to be subjectively worth more than when delayed both by 1 and by 12 months). Such responses are irrational and may result from an erroneous selection of the available alternative. Therefore data of such participants were excluded before any analysis. The percentage of exclusions is quite high. However, in our and other researchers' experience, this percentage is typical of this type of study when conducted online (Appelt et al., 2011; Białek, Markiewicz, & Sawicki, 2015; Białek & Sawicki, 2014). The data reduction criterion used is very sensitive to errors, such that a single incorrect choice affects the whole procedure of calculating the discount rates, and the procedure consists of 36 binary choices which are very similar to each other, potentially resulting in an increased ratio of erroneous choices.

The final sample included 166 participants (91 female) with a mean age of 36.7 years (SD = 10.8, range: 20–69).

Materials and Procedure

Scenario Manipulation

Discount rate calculation was based on the two alternatives forced choice paradigm, in which participants have to choose between smaller sooner (SS) and larger later (LL) alternatives. We have set the discounted amount to $5,000, to reduce the probability of the ceiling effect of discount rates obtained when discounting small gains (Loewenstein & Prelec, 1992), and the delays were set

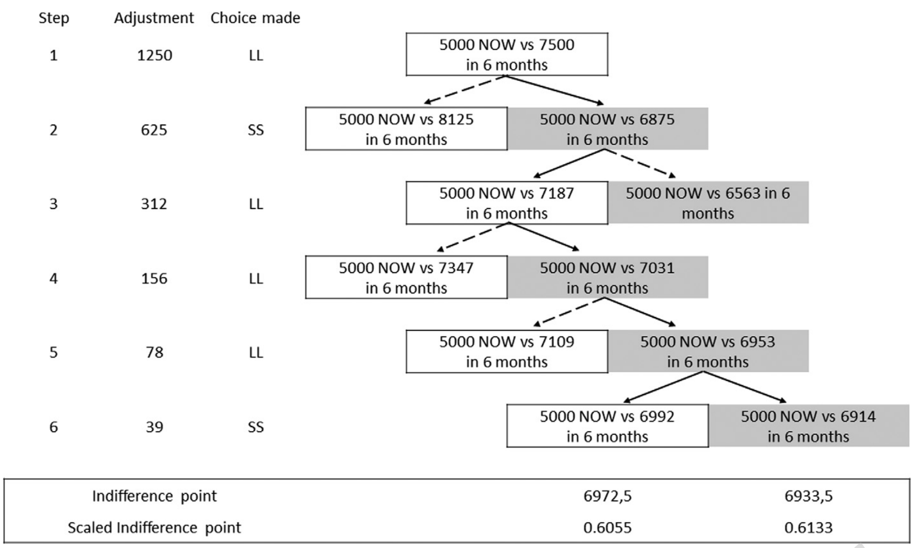

Figure 1. ~~[Author: please provide figure name]~~ An example of amounts presented to a participant in response to his choices in a delaying scenario, with a 6-month delay. Alternatives for selection presented to each participant are shown in boxes. Selection SS results in an increase of LL (gray box) while selecting LL results in its decrease (white box). Solid lines connecting the boxes illustrate problems presented to one of the participants based on his previous choice, while dashed lines show problems that would be presented if the participant had chosen otherwise.

to 1, 6, and 24 months. The exact wording of the instructions is shown in the Electronic Supplementary Material, ESM 1.

Obtaining the Discount Rates: The Adjusting Amount Procedure

After being presented with a scenario, each participant conducted the adjusting amount procedure. Here, in a set of choices, one alternative was fixed while the other was adjusted according to previous choices. Specifically, in the delaying condition, the first step of adjustment involved the choice between fixed SS ($5,000 now) and adjusted LL ($7,500 one month from now). In the accelerating condition, on the other hand, the first step involved the choice between adjusted SS ($2,500 now) and fixed LL ($5,000 one month from now). Although the adjusted alternatives were different in accelerating and delaying conditions ($2,500 or $7,500, respectively) it didn't affect the final discount rates, as the magnitude effect is driven by the discounted amount that was held constant (Sawicki & Białek, 2016).

The direction of the adjustment depended upon the previous choice. When SS was selected, LL would increase to foster competition with the subjective value of SS, whereas selecting LL would decrease its value. The amount of adjustment of the varying alternative decreased by half in each of six consecutive steps, from $1,250 (half of the difference between SS and LL) in the first step to $39 in the sixth step (see Figure 1), after which the procedure stopped. The indifference point between SS and LL is set as a mean of the amounts of the varying values from the two last steps.

In this experiment, SS was fixed at $5,000 for a reference point in the present, and the possible LL equivalent ranged $5,001–$10,000. For a reference point in the future, LL was fixed at $5,000 and SS subsequently adjusted, and the possible SS equivalent ranged $1–$4,999.

Once a participant provided equivalents for each delay of a payoff, the obtained equivalents and delays were standardized and placed in a matrix (Figure 2). Scaling of indifference points differs for delaying and accelerating conditions (Mazur, 1988). For example, supposed an individual has obtained an indifference point for delaying a payoff by 1 month equal to 5,600. His scaled subjective value is (10,000 − 5,600)/5,000 = 4,400/5,000 = 0.88. Now suppose the same person obtained an indifference point for accelerating a payoff equal to 4,750. His scaled subjective value is equal to 4,750/5,000 = 0.95. ~~The points on the matrix were then connected by lines, creating a curve illustrating changes over the delay in the subjective value of payoff.~~ The area under the curve (AUC) was then calculated separately for each participant and used as a measure of discounting strength (Myerson, Green, & Warusawitharana, 2001). We decided to use AUC and not any other method because it is assumption-free and does not require fitting to an exponential or hyperbolic curve. Because of the standardization, the value of the AUC ranged from 0 to 1, with greater values interpreted as weaker discount rates.

Cognitive Reflection

Cognitive reflection was measured using the Cognitive Reflection Test (CRT **[Author: approve abbreviation]**; Frederick, 2005). This test consists of three problems, each having an incorrect but intuitive response and a correct but unintuitive response (see ESM 1 for wording in Polish and English). People who do not simply rely on their first

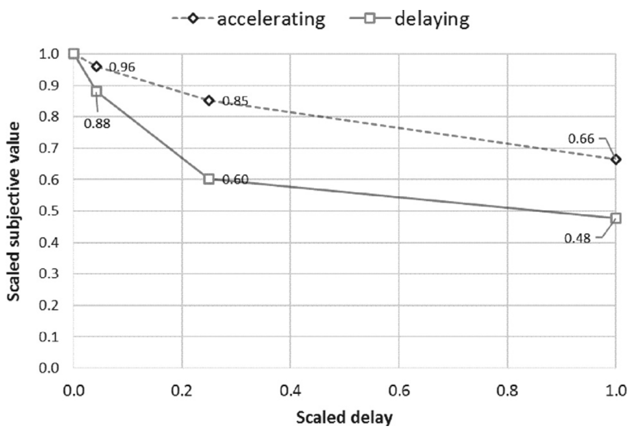

Figure 2. **[Author: please provide figure name]** The area under the curve for delay of a payoff. The indifference points, illustrated by the results obtained from one of the participants, are scaled to the maximum value and placed on the matrix. The starting value is always set to its maximum value, =1, subsequent points have a lower value than the starting value. The area under the curve (AUC) is a sum of trapezoids, created by connecting scaled indifference points and scaled delays, and is interpreted so that greater AUC denotes weaker discounting.

thoughts, but are willing to reflect, tend to score better on the CRT. For example, when considering CRT item no. 2:

> If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?

First response that comes to one's mind is usually "100 minutes." However, this is incorrect, as the number of machines is irrelevant to the amount of time it takes one machine to produce one widget (5 minutes); therefore, any number of machines (e.g., 100) would require only 5 minutes to produce an equivalent number of widgets.

In addition to counting the number of correct responses for each participant, we also controlled for previous CRT experience by directly asking the participants if they had encountered this test before. This was done, as recent research suggests boost of the CRT score in the experienced group (Haigh, 2016; Stieger & Reips, 2016), which can affect the planned tests. Specifically, there is a risk that some part of high CRT scorers is, in fact, low at the cognitive reflection, and can increase the type II error, that is, rejecting true hypothesis.

Results

Table 1 presents the descriptive statistics for discounting strength and reference point dependency, obtained in Experiment 1.

Table 1. Mean discounting strength (AUC) for accelerating and delaying gains depending on individuals' cognitive reflection in Experiment 1

CRT score	*n*	AUC accelerating	AUC delaying	RPDI
0	55	0.68 (0.22)	0.62 (0.25)	0.16 (0.15)
1	29	0.77 (0.16)	0.72 (0.20)	0.15 (0.11)
2	42	0.77 (0.13)	0.72 (0.19)	0.15 (0.15)
3	40	0.78 (0.15)	0.78 (0.15)	0.08 (0.10)
Overall	166	0.75 (0.17)	0.70 (0.21)	0.14 (0.13)

Note. Data presented as mean (standard deviation). CRT = Cognitive Reflection Test; RPDI = Reference Point Dependency Index. **[Author: please approve edit of Note]**

Participants' mean CRT score for the whole sample was 1.40 ($SD = 1.18$). About 40% ($n = 67$) of them declared that they had previously encountered one or more of the CRT tasks, and these participants scored nonsignificantly higher, $M = 1.54$, compared to those with no previous experience, $M = 1.19$, $F(1, 164) = 3.60$, $p = .060$, $\eta^2_p = .021$. This inflated CRT score in the experienced group has been previously found by others. However, it does not affect the CRT's predictive validity (Białek & Pennycook, 2017). Thus, considering recommendations from Białek and Pennycook (2017), we decided to analyze the whole sample combined, using previous experience with the CRT as a covariate for all subsequent analyses involving CRT score.

A one-way within-subjects analysis of variance (ANOVA) confirmed the well-established effect of greater discounting strength for delaying gains compared to accelerating gains, $F(1, 165) = 30.05$, $p < .001$, $\eta^2_p = .154$. There was also a significant but moderate correlation between discounting strength and CRT score in both the delaying condition, Pearson's $r(163) = .28$, $p < .001$, and the accelerating condition, Pearson's $r(163) = .26$, $p < .001$. The correlations do not differ according to their strength (Fisher $Z = .17$, $p = .434$). It should be noted here that greater AUC means weaker discounting. Hence the positive sign of correlation suggests that the increase in CRT is related to weaker discounting. Thus, we replicated evidence for quantitative individual differences between low and high reflective individuals about the discounting strength.

Nonetheless, the study's focus was not on qualitative differences related to individuals' susceptibility to imposed reference points. In short, we wanted to test whether more reflective individuals are less affected by imposed reference points compared to less reflective individuals. To this end, we introduced a new variable – the reference point dependency index (RPDI), using the following formula:

$$\mathrm{RPDI} = \frac{\mathrm{AUC_{acc}} - \mathrm{AUC_{del}}}{(1 - \mathrm{AUC_{acc}}) + (1 - \mathrm{AUC_{del}})}.$$

The assumption behind such formula is that the difference between AUC has to be standardized by the discount rate itself because the same difference between two experimental conditions for impulsive individuals suggests less variability than the same difference for patient individuals.

Thus, RPDI shows whether discounting strength varied according to the participants' reference point. An RPDI close to zero indicates that a person discounts with comparable strength regardless of the reference point, while a nonzero value suggests that the discount strength varies. More specifically, a positive value indicates that discounting strength is greater when delaying than when accelerating, and a negative value shows that discounting strength is greater when accelerating than when delaying.

Because we were not interested in the direction of the change of discount rates but only in its magnitude, we thus used the absolute value of the RPDI for subsequent analysis. To better understand why we decided to use absolute value consider the following: half of the participants discounted more strongly in accelerating and another half in delaying condition. The overall effect of reference point would be zero, although none of the participants displayed consistent discount rates. Using the absolute value solves this issue, and allows to distinguish the bidirectional effect of reference points from independence from reference points.

Participants' raw CRT scores were regressed on the absolute value of the RPDI controlling for participants' previous CRT experience (the effect of previous experience was nonsignificant, p = .828). CRT score significantly predicted the RPDI, $\beta = -.20$, $t(163) = -2.58$, $p = .011$. The negative value of β suggests that greater CRT scores predicted greater resistance to externally imposed reference points.

Discussion

Although our data suggest that cognitively reflective individuals are less impulsive and less affected by imposed reference points, there are several motives to conduct Experiment 2. First of all, the quality of our sample is arguably not ideal because the participants conducted the study online, and we had to exclude a nontrivial percentage of participants. Here is a possibility, that majority of excluded individuals are in fact also a low cognitively reflective sample. Speculatively, the level of exclusions can be reduced in the laboratory, where experimenter could infrequently remind participants to "try-hard," boosting this way not only the compliance with instructions, but also the psychometric characteristics of the tests (Steinborn, Langner, & Huestegge, 2017; but see Hedge, Powell, & Sumner, 2017 ~~[Author: add to references]~~ for discussion on how high-reliability paradox is studied on individual differences). However, to avoid introducing new factors that could affect results of Experiment 2 and impede comparison between Experiments 1 and 2, we refrained from doing so.

Another issue is that cognitive reflection correlates, among other things, with cognitive abilities (Toplak, West, & Stanovich, 2014) and numeracy (Sinayev & Peters, 2015). Both these variables are also involved in financial decision-making. Thus, the correlation between CRT and the strength of discounting can be artificially shaped by their mutual correlations with numeracy (for discussion see Sinayev & Peters, 2015).

Experiment 2

In this experiment, we aimed to replicate the findings of Experiment 1, with additional control of individuals' numeracy. Also, we conducted this study in a laboratory environment in the hopes of reducing the number of nonsystematic discounting data obtained in Experiment 1.

Participants

One hundred and thirteen Polish speaking participants recruited via Online Recruitment Software for Economic Experiments (ORSEE) participated in the experiment in exchange for $8. From this sample, 13 participants (11.5% of the sample) have been removed because of nonsystematic discounting data. From the remaining $n = 100$ sample, 40 were female, and their average age was 34.8 years (SD = 17.9, range: 19–74). Because of the coding error, providing age was not mandatory, thus this estimate is based on the data of only 64 individuals.

Materials and Procedure

The first two steps of the experimental procedure (measurement of discounting and cognitive reflection) remained unchanged from Experiment 1, with a small difference of the language it was administered (Polish vs. English) and the amounts discounted (5,000 PLN vs. $5,000). The change in discounted amount shouldn't affect the discount rates, as the values have a similar relation to average monthly gross salary in both countries and are numerically identical.

In addition to the tasks used in Experiment 1, the Berlin Numeracy Test (BNT; Cokely, Galesic, Schulz, Ghazal, & Garcia-Retamero, 2012) was administered. This short task captures individual numeracy skill via four short mathematical tasks, such as:

> Imagine we are throwing a five-sided die 50 times. On average, out of these 50 throws how many times would this five-sided die show an odd number (1, 3 or 5)?

The difference between the BNT and the CRT is that the latter, but not the former, has a strong intuitively accessible invalid response. Hence, numerical skills required in BNT are not sufficient for CRT, as it additionally requires one to override the intuitive response with the less salient, but correct response. Hence, using both, BNT and CRT, allows evaluating the contribution of numeracy and reflection to delay discounting. We also collected information on participants' previous exposure to the corresponding tests.

The wording of all tasks used in Experiment 2 can be found in ESM 1.

Table 2. Mean discounting strength (AUC) for accelerating and delaying gains depending on individuals' cognitive reflection in Experiment 2

CRT score	*n*	AUC accelerating	AUC delaying	RPDI
0	27	0.73 (0.15)	0.50 (0.29)	0.30 (0.19)
1	16	0.74 (0.23)	0.51 (0.30)	0.34 (0.26)
2	29	0.67 (0.25)	0.52 (0.31)	0.21 (0.13)
3	28	0.78 (0.14)	0.66 (0.18)	0.24 (0.15)
Overall	100	0.73 (0.20)	0.55 (0.28)	0.27 (0.18)

Note. Data presented as mean (standard deviation). CRT = Cognitive Reflection Test; RPDI = Reference Point Dependency Index. [Author: please approve edit of Note]

Results

Participants scored on average 1.58 (SD = 1.17) in the CRT, and 1.48 (SD = 1.21) in the BNT. Consistent with Experiment 1, individuals having experience with the CRT (n = 28, m = 2.25) scored higher than inexperienced individuals (n = 72, m = 1.32), $F(1, 99) = 14.64$, $p < .001$, $\eta^2_p = .130$. There were only four individuals who previously encountered the BNT. Hence no statistical tests were calculated.

Table 2 presents the descriptive statistics for discounting strength and reference point dependency with respect to the CRT score, obtained in Experiment 2.

As in the previous experiment, we replicated the direction effect of stronger discounting of delaying versus accelerating gains, $F(1, 99) = 86.62$, $p < .001$, $\eta^2_p = .467$. We also tested the correlation between CRT scores and discounting strength, controlling for CRT experience, this time finding a significant correlation with AUC_{del}, $r(98) = .251$, $p = .012$, but not with AUC_{acc}, $r(98) = .064$, $p = .529$. BNT scores did not correlate with any discounting parameters, $p > .300$, and BNT and CRT scores correlated strongly, $r(98) = .552$, $p < .001$.

Finally, raw scores of CRT, BNT, and their interaction product were regressed on the RPDI, with CRT experience held as a covariate (which didn't predict the RPDI, $p = .193$). Results showed significant effect of the CRT, $\beta = -.40$, $t(95) = -2.39$, $p = .019$, but not of the BNT, $p = .206$, nor its interaction, $p = .134$.

To sum up, the findings of Experiment 2 are fully consistent with the main findings of Experiment 1.

General Discussion

We found two types of individual differences related to delay discounting. First, the quantitative difference is that more reflective individuals tend to be less impulsive than less reflective individuals. Second, the qualitative difference is that such individuals are more resilient to imposed reference points than less reflective individuals. Hence, discounting of cognitively reflective individuals was weaker and less dependent on which alternative was default.

We see two possible explanations for this result: (1) reflective individuals might cope better with the greater number of impulsive thoughts that arise in a delaying condition compared to an accelerating condition (a resisting impulsive thoughts strategy); and/or (2) more reflective individuals might collect and consider more information before making a decision (an information inquisition strategy).

The argument for resisting impulsive thoughts strategy suggests that people process information regardless of their cognitive reflection, but their ability to control their judgments and decisions helps them to cope better with impulsive thoughts (Toplak et al., 2011). Thus, even when the number of impulsive thoughts is greater, as it happens when the immediate reward is the reference point, their judgments seem to be similar to judgments constructed when its number is low.

A potential concern regarding this explanation is that impulsive thoughts are more vivid in the delaying scenario when the reference point is in the present. Hence, more reflective individuals, who resist impulsive thoughts better, should differ from less reflective individuals mostly in their delaying discount rates. The rationale here is simple: you should see a greater impact of reduced influence of impulsive thoughts when there are more of such thoughts.

We found mixed support for this claim, as in Experiment 1 the CRT predicted discounting strength equally well in accelerating and delaying conditions, while in Experiment 2, consistently with the predictions of resisting impulsive thoughts strategy, CRT predicted the discounting strength only in the delaying condition.

The resisting impulsive thoughts explanation would predict weaker discounting strength, thus smaller differences between discount rates, as subtracting small numbers is likely to produce small differences. Even when controlling for the impulsivity *per se*, that is, scaling the difference between discount rates by the strength of discounting, more reflective individuals were more likely to show the same discount rates regardless of the imposed reference point.

Hence, we suggest the information inquisition strategy as the more promising explanation where less reflective individuals tend to make decisions early, as soon as they considered the first alternative and more reflective people tend to delay their decisions until after they have considered both choice alternatives. Hence, cognitive reflection might decrease the impact of the order of thoughts, which means that less reflective individuals would care more for thoughts coming first to their mind and less for subsequent thoughts, while more reflective individuals would weight all thoughts more equally. Thus, judgments of reflective individuals, as long as the reference points are irrelevant to the core problem, can be described as more efficient and rational.

Notably, we provided evidence that cognitive reflection predicts individual differences in delay discounting over and above individuals ability to process numerical information, as captured by the BNT. Hence, we can refute the explanation suggested by some researchers that effects of cognitive reflection can be explained merely by better numeracy of cognitively reflective individuals.

Future studies could shed more light on the relationship between cognitive reflection and delay discounting, for example, by testing the reported effects both in the domain of gains and losses. It could be useful to control for participants engagement in the task and associated affective factors such as worry, distress using, for example, Dundee Stress State Questionnaire (DSSQ; Langner, Steinborn, Chatterjee, Sturm, & Willmes, 2010; Matthews et al., 1999).

Cognitive reflection can be understood as a motivational feature of cognition (Pennycook, Fugelsang, & Koehler, 2015; Pennycook & Ross, 2016). Hence, boosting one's willingness to engage in reflection could decrease the impact of frames and reference points on intertemporal choices, improving the consistency and rationality of judgment and decision-making. Manipulations on cognitive reflection should, therefore, be used to improve the decision-making of individuals.

Electronic Supplementary Material

The electronic supplementary material is available with the online version of the article at https://doi.org/10.1027/1614-0001/a000254

ESM 1. Text (doc).
[Author: Please provide a short description of the content of ESM]

References

Appelt, K. C., Hardisty, D. J., & Weber, E. U. (2011). Asymmetric discounting of gains and losses: A query theory account. *Journal of Risk and Uncertainty, 43*, 107–126. https://doi.org/10.1007/s11166-011-9125-1

Baron, J. (2016). A welfarist approach to manipulation. *Journal of Marketing Behavior, 1*, 283–291. https://doi.org/10.1561/107.00000018

Benzion, U., Rapoport, A., & Yagil, J. (1989). Discount rates inferred from decisions: An experimental study. *Management Science, 35*, 270–284. https://doi.org/10.1287/mnsc.35.3.270

Białek, M., Markiewicz, Ł., & Sawicki, P. (2015). Introducing conjoint analysis method into delayed lotteries studies: Its validity and time stability are higher than in adjusting. *Frontiers in Psychology, 6*, 23. https://doi.org/10.3389/fpsyg.2015.00023

Białek, M., & Pennycook, G. (2017). The Cognitive Reflection Test is robust to multiple exposures. *Behavior Research Methods*. Advance online publication. https://doi.org/10.3758/s13428-017-0963-x

Białek, M., & Sawicki, P. (2014). Can taking the perspective of an expert debias human decisions? The case of risky and delayed gains. *Frontiers in Psychology, 5*, 989. https://doi.org/10.3389/fpsyg.2014.00989

Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring risk literacy: The Berlin Numeracy Test. *Judgment and Decision Making* **[Author: add doi]**, *7*, 25–47.

De Neys, W. (2014). Conflict detection, dual processes, and logical intuitions: Some clarifications. *Thinking & Reasoning, 20*, 169–187. https://doi.org/10.1080/13546783.2013.854725

De Neys, W., & Bonnefon, J.-F. (2013). The "whys" and "whens" of individual differences in thinking biases. *Trends in Cognitive Sciences, 17*, 172–178. https://doi.org/10.1016/j.tics.2013.02.001

Frederick, S. (2005). Cognitive reflection and decision making. *The Journal of Economic Perspectives, 19*, 25–42. https://doi.org/10.1257/089533005775196732

Haigh, M. (2016). Has the standard Cognitive Reflection Test become a victim of its own success? *Advances in Cognitive Psychology, 12*, 145–149. https://doi.org/10.5709/acp-0193-5

Hardisty, D. J., Thompson, K. F., Krantz, D. H., & Weber, E. U. (2013). How to measure time preferences: An experimental comparison of three methods. *Judgment and Decision Making* **[Author: add doi]**, *8*, 236–249.

Johnson, E. J., Häubl, G., & Keinan, A. (2007). Aspects of endowment: A query theory of value construction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 461–474. https://doi.org/10.1037/0278-7393.33.3.461

Langner, R., Steinborn, M. B., Chatterjee, A., Sturm, W., & Willmes, K. (2010). Mental fatigue and temporal preparation in simple reaction-time performance. *Acta Psychologica, 133*, 64–72. https://doi.org/10.1016/j.actpsy.2009.10.001

Lee, L., Lee, M. P., Bertini, M., Zauberman, G., & Ariely, D. (2015). Money, time, and the stability of consumer preferences. *Journal of Marketing Research, 52*, 184–199. https://doi.org/10.1509/jmr.10.0386

Loewenstein, G., & Prelec, D. (1992). Anomalies in intertemporal choice: Evidence and an interpretation. *The Quarterly Journal of Economics, 107*, 573–597. https://doi.org/10.2307/2118482

MacKillop, J., Amlung, M. T., Few, L. R., Ray, L. A., Sweet, L. H., & Munafò, M. R. (2011). Delayed reward discounting and addictive behavior: A meta-analysis. *Psychopharmacology, 216*, 305–321. https://doi.org/10.1007/s00213-011-2229-0

Matthews, G., Joyner, L., Gilliland, K., Campbell, S., Falconer, S., & Huggins, J. (1999). Validation of a comprehensive stress state questionnaire: Towards a state big three. In I. I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in europe* (7th ed., pp. 335–350). Tilburg, The Netherlands: Tilburg University Press.

Mazur, J. E. (1988). Estimation of indifference points with an adjusting-delay procedure. *Journal of the Experimental Analysis of Behavior, 49*, 37–47. https://doi.org/10.1901/jeab.1988.49-37

Myerson, J., Green, L., & Warusawitharana, M. (2001). Area under the curve as a measure of discounting. *Journal of the Experimental Analysis of Behavior, 76*, 235–243. https://doi.org/10.1901/jeab.2001.76-235

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015). What makes us think? A three-stage dual-process model of analytic engagement. *Cognitive Psychology, 80*, 34–72. https://doi.org/10.1016/j.cogpsych.2015.05.001

Pennycook, G., & Ross, R. M. (2016). Commentary: Cognitive reflection vs. calculation in decision making. *Frontiers in Psychology, 7*, 9. https://doi.org/10.3389/fpsyg.2016.00009

Sawicki, P., & Białek, M. (2016). Side effects in time discounting procedures: Fixed alternatives become the reference point. *PLoS One, 11*, e0165245. https://doi.org/10.1371/journal.pone.0165245

Sinayev, A., & Peters, E. (2015). Cognitive reflection vs. calculation in decision making. *Frontiers in Psychology, 6*, 532. https://doi.org/10.3389/fpsyg.2015.00532

~~Stanovich, K. E., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology* **[Author: add in-text citation or delete reference]**, *94*, 672. https://doi.org/10.1037/0022-3514.94.4.672~~

Steinborn, M. B., Langner, R., & Huestegge, L. (2017). Mobilizing cognition for speeded action: Try-harder instructions promote motivated readiness in the constant-foreperiod paradigm. *Psychological Research, 81*, 1135–1151. https://doi.org/10.1007/s00426-016-0810-1

Stieger, S., & Reips, U.-D. (2016). A limitation of the Cognitive Reflection Test: Familiarity. *PeerJ, 4*, e2395. https://doi.org/10.7717/peerj.2395

Thaler, R. H., & Shefrin, H. M. (1981). An economic theory of self-control. *Journal of Political Economy, 89*, 392–406. https://doi.org/10.1086/260971

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition, 39*, 1275–1289. https://doi.org/10.3758/s13421-011-0104-1

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking & Reasoning, 20*, 147–168. https://doi.org/10.1080/13546783.2013.844729

Weber, E. U., Johnson, E. J., Milch, K. F., Chang, H., Brodscholl, J. C., & Goldstein, D. G. (2007). Asymmetric discounting in intertemporal choice a query-theory account. *Psychological Science, 18*, 516–523. https://doi.org/10.1111/j.1467-9280.2007.01932.x

~~West, R. F., Toplak, M. E., & Stanovich, K. E. (2008). Heuristics and biases as measures of critical thinking: Associations with cognitive ability and thinking dispositions **[Author: add in-text citation or delete reference]**. *Journal of Educational Psychology, 100*, 930–941. https://doi.org/10.1037/a0012842~~

Received February 28, 2017
Revision received August 28, 2017
Accepted October 11, 2017
Published online XX, 2018

Michał Białek
Centre for Economic Psychology and Decision Sciences
Koźminski University
Jagiellońska 59
03-301 Warsaw
Poland
mbialek@kozminski.edu.pl