Cognitive Reflection But Not Reinforcement Sensitivity Is Consistently Associated With Delay Discounting of Gains and Losses

Michał Białek
University of Waterloo and Kozminski University

Max Bergelt
University of Waterloo

Yoshimasa Majima
Hokusei Gakuen University

Derek J. Koehler
University of Waterloo

We investigated potential predictors of the preference for smaller sooner over larger later monetary payoff, so-called delay discounting. Cognitive reflection is the tendency to override intuitive responses with deliberative ones. Reinforcement sensitivity captures the strength of an individual's reaction to possible gains or losses in their environment. In 2 studies ($n = 249$), 1 exploratory and 1 preregistered, we found that cognitive reflection, but not reinforcement sensitivity, consistently predicted temporal discounting of hypothetical monetary gains and losses. We conclude that discount rates depend on how people combine various thoughts rather than on how strong is their sensitivity to gains and losses.

Keywords: discounting, impatience, reinforcement sensitivity, cognitive reflection, numeracy

Individuals who fail to delay gratification show increased risk of various addictions (Bickel, Odum, & Madden, 1999; MacKillop et al., 2011; Perry, Larson, German, Madden, &

Michał Białek, Department of Psychology, University of Waterloo, and Department of Economic Psychology, Kozminski University; Max Bergelt, Department of Psychology, University of Waterloo; Yoshimasa Majima, Department of Psychology for Well-Being, Hokusei Gakuen University; Derek J. Koehler, Department of Psychology, University of Waterloo.

Correspondence concerning this article should be addressed to Michał Białek, who is now at Institute of Psychology, University of Wrocław, Dawida 1, 50-529 Wrocław, Poland. E-mail: michal.bialek3@uwr.edu.pl

Carroll, 2005), health issues (Daugherty & Brase, 2010), criminal and antisocial behavior (Mishra & Lalumiere, 2017), and lower socioeconomic status (Moffitt et al., 2011). The extent to which people fail to delay payoffs is captured by discount rates. Such intertemporal choices are malleable, affected by contextual factors, for example, the reference point in the presence increases discount rates more than the reference point in the future (Appelt, Hardisty, & Weber, 2011; Loewenstein, 1988). Another example is that as early as in 4-year-old children, an unreliable environment promotes immediate gratification compared with maximization of rewards in a reliable environment (Kidd, Palmeri, & Aslin, 2013).

Neuroscientific research has identified certain areas of the brain that are associated with delay discounting (Peters & Büchel, 2011; Shamosh et al., 2008). Discounting appears to involve a number of distinct neural systems including a valuation system that represents the incentive value of a reinforcer (regions in the orbitofrontal

cortex, the ventral striatum, and the posterior cingulate cortex), a cognitive control system that deals with decision conflict (regions in the lateral prefrontal cortex and the anterior cingulate cortex) and integration of information (anterior prefrontal cortex), and a predictive/affective system that represents decision outcomes (regions in the medial temporal lobes such as the hippocampus and amygdala).

The question is what are the reasons why people differ in their discount rates. It could be that they experience different input from the external world (i.e., are differentially sensitive to immediate rewards and punishments) or that they process information differently (e.g., tend to rely on more vivid information when forming their preferences). In other words, it is critical to understand whether people who fail to delay gratification and extensively delay negative outcomes are simply experiencing stronger temptations or whether they integrate reward and delay information differently.

We broadly classify the predictors of delay discounting into two categories. The first group of possible predictors relates to the concept of reinforcement sensitivity: the extent to which an organism reacts to possible payoffs or losses in their environment. This group thinks of discounting as a difference in how the external world is internally represented, that is, that an individual who experiences immediate gains and losses more strongly is more inclined to immediately obtain gains and to postpone losses. According to reinforcement sensitivity theory, attitudes toward gains and losses are the main processes underlying the development of personality (Corr, 2004; Gray & McNaughton, 2003). The most widespread version of this theory proposes two independent systems: a behavioral activation system (BAS) that responds to conditioned and unconditioned reward cues and a behavioral inhibition system (BIS) that responds to punishment and nonrewarding stimuli (Gray, 1982). Although this model required some revisions (Gray & McNaughton, 2003), the idea that our reactions to reinforcement are a major predictor of our behavior, remains unchanged.

The relationship between reinforcement sensitivity and discounting has some empirical support, for example, activation in reward-related regions of the brain was correlated with stronger delay discounting (Hansen, Turpyn, Mauro, Thompson, & Chaplin, 2019), and Behavioral Activation Scale subscales like drive and reward-responsiveness correlated positively with Behavioral Impulsivity measure, where in 50 trials participants could receive 5 cents after a brief period of waiting versus 15 cents for longer periods of waiting (Corvi, Juergensen, Weaver, & Demaree, 2012). A recent study in bipolar disorder and alcohol dependent individuals showed that higher reward sensitivity correlated with weaker delay discounting (Mellick, Tolliver, Brenner, & Prisciandaro, 2019). Furthermore, stronger discounting of monetary reward was predicted by higher scores in the fun-seeking subscale of the BAS (Jarmolowicz et al., 2014), or in sensitivity to rewards, but only after being primed by sexy commercials (Van den Bergh, Dewitte, & Warlop, 2008). Finally, effort discounting was only predicted by the drive subscale of the BAS (Siegelman, 2016). On the other hand, individuals who anticipated greater rewards in the future tended to discount less strongly, evidencing that greater sensitivity toward future reward can predict weaker discounting (Benningfield et al., 2014), and those who showed diminished neural responses to anticipated rewards at the age 14 were also problematic drug users at the age of 16 (Büchel et al., 2017).

The sign of the reinforcement sensitivity correlation with delay discounting is therefore an open question. In light of the two latter studies, reinforcement sensitivity should predict a drive toward the biggest reward or smallest loss; in the light of all other studies, reinforcement sensitivity should predict a drive toward immediate gains and toward most delayed losses. What can be predicted, however, is that there should be a correlation between the two variables.

The second group of possible predictors focuses on the way people process information, where people make internal trade-offs between several goals that they can satisfy, meaning every decision requires the construction of new preferences toward choice options (Bettman, Luce, & Payne, 1998). This process of constructing preferences is malleable. For example, people tend to be affected by the imposed reference point for their choice: how much a premium one wants to delay immediate gain or how much one is willing to pay for accelerating its receipt (Loewenstein, 1988). Similarly, asking people to consider options in a particular

order (i.e., smaller sooner vs. larger later first) affects their discount rates (Appelt et al., 2011; Sawicki & Białek, 2016, 2017; Weber et al., 2007).

The way people judge and decide is thought of as a function of task properties and individual differences in cognitive style and cognitive abilities. Illustrative evidence comes from studies showing that people who show higher cognitive reflection, and override their first thoughts with reflection (Białek & Sawicki, 2018; Frederick, 2005; Hardisty & Weber, 2009; McClure, Laibson, Loewenstein, & Cohen, 2004; Noori, 2016; Stanovich & West, 2008), and those whose reflection is more efficient because of higher intelligence (Shamosh & Gray, 2008) or greater working memory capacity (Hirsh, Morisano, & Peterson, 2008; Shamosh et al., 2008) tend to discount less strongly. This implies that going for smaller sooner gains is intuitive. To select larger later gains, one needs to reflectively override this intuition. In addition, the effects of reference points seem to be mitigated by cognitive reflection, where discount rates are less dependent on which of the two choice options were presented to participants as the reference point (Białek & Sawicki, 2018). Therefore, cognitive style and cognitive ability may contribute to how people combine external cues and their goals in creating their intertemporal preferences, with higher scores on these traits leading to choices that are resistant to irrelevant contextual cues.

We consider cognitive reflection to be the most promising predictor of delay discounting. This is because cognitive reflection is sometimes labeled as willingness to reflect, which assumes some level of conscious decision whether to reflect or not, and therefore is a precondition of higher-order thinking (Pennycook & Ross, 2016; Toplak, West, & Stanovich, 2014). Although working memory capacity, cognitive abilities, and cognitive reflection are all correlated, people who fail to engage in reflection will not make use of their cognitive abilities to override intuitive responses (Pennycook, Fugelsang, & Koehler, 2015; Stanovich, 2018; Stanovich & West, 2008), making cognitive reflection a precondition of making any use of their cognitive abilities. One of the explanations of why cognitive reflection is linked to delay discounting is that it allows individuals to override their intuitive and impulsive thoughts

through reflection or to integrate multiple goals, any of which promotes weaker delay discounting. This position suggests a qualitative difference between individuals with greater and smaller discount rates, in which the decisive factor is whether or not people override their impulsive thoughts. To remind you, reinforcement sensitivity would link the strength of delay discounting with the strength of the impulsive thoughts.

Finally, both positions may be true as reinforcement sensitivity might be linked to cognitive reflection. This would mean that more reflective individuals tend to be less affected by reinforcement and thus discount less steeply or that such individuals cope better with weaker impulsive thoughts.

To understand discounting behavior, we investigated hypothetical financial choices for both gains and losses. This distinction between delay discounting behavior in gains and in losses is critical, as these demonstrate some quantitative difference (the so-called sign effect; Chapman, 1996; Thaler, 1981), whereby rewards are discounted at higher rates than gains are. Discounting of gains and losses also demonstrates some qualitative differences (Abdellaoui, Attema, & Bleichordt, 2010; Estle, Green, Myerson, & Holt, 2006); for example, changing reference points affects discounting of gains but not of losses (Sawicki & Białek, 2017), and the size of a loss has only a minor impact, whereas the size of a gain has a major impact on discount rates (Estle et al., 2006; Mitchell & Wilson, 2010). Finally, discounting of losses can be also affected by dread avoidance: people's desire to get anticipated losses over with as fast as possible (Loewenstein, 1987; Story et al., 2013), even disregarding the utility of the considered outcomes (Benhabib, Bisin, & Schotter, 2010).

To our knowledge, our project is the first attempt at testing reinforcement sensitivity and cognitive style jointly in the context of intertemporal choice. Not only did previous work test these effects separately, it also tended to focus solely on the discounting of gains (Corvi et al., 2012; Hansen et al., 2019). In testing for predictors of both gains and losses, we make a valuable contribution to the understanding of delay discounting.

Testing for cognitive reflection usually requires participants to answer several mathematical questions with intuitive and incorrect re-

sponses. Therefore, numeracy contributes to performance on the test but is not conceptualized as a component of cognitive reflection (Campitelli & Gerrans, 2014; Sinayev & Peters, 2015; Sirota, Kostovičová, Juanchich, Dewberry, & Marshall, 2018). Specifically, one can perform poorly on the cognitive reflection test not because one fails to engage in reflection but because one simply failed at completing the required calculations. Therefore, numeracy can be a confounding factor when evaluating participants' cognitive reflection, or it even explain the effects of the Cognitive Reflection Test (CRT) completely. Because of this, we decided to add a numeracy test to better assess the unique contribution of cognitive reflection in explaining delay discounting. For illustration, in our related work (Białek & Sawicki, 2018), performance on the CRT predicted consistency in delay discounting across varying reference points with $\beta = -.20$. But when we controlled for numeracy, the predictive power of the CRT doubled, to $\beta = -.40$, whereas numeracy was not a significant independent predictor. Considering the above, to identify potential predictors of discount rates, we tested for participants' reinforcement sensitivity and for cognitive reflection while also controlling for numeracy.

Experiment 1

Participants

A total of 200 participants were recruited via TurkPrime (Litman, Robinson, & Abberbock, 2017), and they participated in exchange for $1. Participants had to have previously completed at least 100 studies with a 90% acceptance rate and were only from the United States. It is a standard practice in delay discounting research to eliminate participants with zero or negative discount rates because they represent qualitatively different cognitive process. Excluding such data is the recommendation from Yoon and Chapman (2016), who designed the Three-Option Adaptive Discount Rate Measure (ToAD), which simply returns no results for participants who displayed negative discount rates. Considering the above, we excluded data from 30 participants because their discount rates were either zero or negative and from a further 46 who failed to respond to all items.[1,2] For this experiment, we analyzed data from the remain-

ing $n = 124$ participants (55 female, 70 male, one other, $M_{age} = 37.8$, range: 21–74, $SD = 12.7$). Although intended to be bigger, our sample size was still large enough to detect correlations as small as $r = .22$ with 0.8 power. We consider smaller correlations as unlikely to substantially contribute to explaining the discount rates. Data for this and the second experiment are available at https://osf.io/8k69y/?view_only=ba303396b77a466bbbae682ffc82da2b.

Method

For this research we used four questionnaires and one experimental task. We briefly describe these tools in the following text.

Reward Responsiveness Questionnaire. The Reward Responsiveness Questionnaire (RRQ) was developed by Van den Berg, Franken, and Muris (2010) and was designed to measure peoples' differing responsivity to rewards. The questionnaire consists of eight items, with each item being a statement that a person may agree or disagree with. The items are measured on a 4-point scale from 1 (*strong disagreement*) to 4 (*strong agreement*). An example of an item used is "When I'm doing well at something, I love to keep at it."

Behavioral Inhibition System Questionnaire. The Behavioral Inhibition System Questionnaire (BISQ) was taken from the BIS/BAS scales developed by Carver and White (1994), which consists of 24 items assessing individual variability in the sensitivity of two motivational systems: the BIS and the BAS. The BISQ consists only of the seven items that assess the BIS. Each item is a statement that may either be true or false of a person. The items are measured on a 4-point scale from 1 (*very true for me*) to 4

---

[1] However, for exploratory reasons, we replaced missing data with zero discount rates and recalculated all our findings. The results are fully consistent with the ones reported here.

[2] Consistent with ethics guidelines of the University of Waterloo, we allowed participants to skip questions and still be eligible for compensation. The number of participants who failed to respond to all items suggests that our Mechanical Turk participants promptly used this to gain their payoff without spending too much time on the task. Alternatively, the high rates of omissions in Experiments 1 and 2 was caused by lack of any reminder of missing responses, which otherwise would prevent participants from skipping items.

(*very false for me*). An example of an item used is "Criticism or scolding hurts me quite a bit."

Cognitive Reflection Test. The original three-item cognitive reflection test was introduced by Frederick (2005) and was designed to assess a person's tendency to suppress initial intuitive answers that come to mind when presented with certain questions. There are several parallel versions of the CRT, and we decided to use two of the newly developed ones. We did this due to the popularity of the original test and because there is a concern that many participants will have been exposed to the test items (Haigh, 2016). In this study, we incorporated questions from CRT versions developed by Toplak et al. (2014) and by Primi, Morsanyi, Chiesi, Donati, and Hamilton (2016). These two versions of the CRT share one item, and therefore the test we administered had only six items. The correct responses to the test items were combined, creating the "composite" score. Answers to five of these six items were open ended, with one of the items from Toplak et al. (2014) being multiple choice.

All the items in the CRT have a correct answer as well as an intuitive incorrect answer. For example, for one of the items used, "Jerry received both the 15th highest and the 15th lowest mark in the class. How many students are there in the class?", the correct answer is 29, whereas the initial intuitive answer is 30.

Berlin Numeracy Test. The Berlin Numeracy Test (BNT; Cokely, Galesic, Schulz, Ghazal, & Garcia-Retamero, 2012) measures a person's statistical numeracy and probability literacy. The test used was the "traditional paper and pencil format" from the Cokely et al. article, composed of four open-ended items that all involve probability calculations. An example of an item used is "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)?" In this case the correct answer is 30, and as opposed to the CRT, there would be no intuitive but incorrect answer.

Discount rate measure (ToAD). The ToAD was developed by (Yoon & Chapman, 2016) to measure peoples' varying rates of temporal discounting. ToAD is an adaptive protocol that generates forced choice questions with three options. These choices used delays in the range from 0 to 364 days and were set so that the

mean discounted value was $5,000. For example, first a participant was presented with $1,002.81 today, $3,550.07 in 116 days, and $5,227.15 in 224 days. The participant selects which option they would prefer. After this initial step, a discount factor k was broadly estimated, and used to inform choice options presented in the next step. For example, if the participant selected the immediate option as most preferred (i.e., $1,002.81 today, already showing signs of strong discounting), the three new choice options were $242.23 today, $2,770.17 in 268 days, and $4,793.48 in 510 days. Let us go with the immediate payoff once again. After the second step, an even better proxy of k was calculated, and three further options were generated: $113.69 today, $3,338.94 in 120 days, and $5,348.58 in 200 days. Note how the size of the immediate gain decreases, whereas the size of alternative payoffs remains similarly delayed and of similar magnitude. This way, the options are more and more similar in terms of their attractiveness to the participant. This procedure is repeated 10 times, to arrive at a well-calibrated discount rate.[3]

Subjective value of a payoff decreases as its delay increases, with this having a nonlinear shape; that is, the decreases are greatest in the first periods of delay (Mazur, 1987). For example, an immediate gain of $100 is subjectively worth $85 obtained in a month and $70 obtained in a year. The best fit to such a curve is given by a hyperbolic function: $f(x) = 1/(1 + kD)$, where k is the hyperbolic discount rate and D is the delay. Participants' responses were fitted to such hyperbolic curve, where parameter k corresponds to its steepness, interpreted as strength of discounting. In the current work, we adapted

---

[3] Contrary to past research, we could not compare discount rates in gains and losses. This is because of an issue with ToAD identified by David Hardisty (personal communication), whereby the dynamic adjustment procedure confounds magnitude with sign. For example, if the ToAD magnitude is set to $5,000, then the gain version will set larger later amounts equal to $5,000 on every trial, and the smaller sooner amounts and middle amounts will be lower (like, $3,000 today). The loss version will set SS equal to $5,000, and then the larger later amounts and middle amounts will be larger (for example, $7,000 in 180 days). Hence, discounted values are larger in losses than in gains, which in turn has its own effects on discount rates. For future research, researchers should use a new, fixed by the original author, ToAD.

the procedure suggested by the original authors of the ToAD and log-transformed the k value (log k). Here, higher log k values should be understood to represent a greater discounting strength. By log-transforming the k parameter, we reduced the skew of the empirically obtained distribution, which helps in interpreting patterns in data. This is often required in studies on intertemporal choice, as the distribution of empirical discount rates is right skewed, with outliers who express extremely high discount rates, and a majority who express small to medium discount rates.

Procedure

Participants responded to the RRQ, BISQ, CRT, BNT, ToAD gains, and ToAD losses in an online survey run in Qualtrics. These materials were delivered in the fixed sequential order mentioned earlier. Participants were not able to return to previously answered sections of the survey.

Results

To test all our hypotheses, we used Bayesian analysis. Bayesian statistics relies on the same computations as the classically used frequentist statistics, but instead of returning a p value for every test, it returns a Bayes factor. The Bayes factor allows the quantification of evidence for and against a particular hypothesis. For example, $B_{10} = 6$ suggests that obtaining the results found is six times more likely under a tested hypothesis than under the null hypothesis. Conversely, $B_{01} = 6$ suggests that obtaining the results found is six times more likely under the

null hypothesis than under a tested hypothesis. Usually, $B \approx 3$ corresponds to a p value of .05 (Held & Ott, 2018). For more complex comparisons, including more than merely two models, we reported B_M, which represents how much better a given model is compared with the set of all other models. This is followed up by direct comparisons between this model and the next best models, with B_{01} used to describe the evidence for the model alternative to the best candidate. Before testing our hypotheses, we confirmed that the internal reliability of our measures was adequate by calculating the Cronbach's α for all the scales used in this study and correlated all variables in order to assess their interrelations (Table 1).

Consistent with reinforcement sensitivity theory, the BISQ and RRQ scales were independent (Corr, 2004; Van den Berg et al., 2010). Confirming previous findings, cognitive reflection was positively correlated with numeracy (Liberali, Reyna, Furlan, Stein, & Pardo, 2012) and negatively correlated with discounting strength of gains (Białek & Sawicki, 2018; Noori, 2016). The high internal consistency of our materials, close to values reported in the original articles, jointly with replication of previous findings, assures us that our data are valid and reliable despite the substantial dropout rate in our sample.

Next, we conducted two separate Bayesian regression analyses to estimate how the RRQ, BISQ, and CRT predicted discounting strength of gains and losses. Numeracy was excluded from this analysis because it correlated too strongly with cognitive reflection. Figure 1 (Panel A gains, Panel B losses) presents mar-

Table 1
Correlations Between Reinforcement Sensitivity, Cognitive Reflection, and Delay Discounting in Experiment 1

Dependent Variable	M	SD	1	2	3	4	5	6
1) Reward responsiveness (RRQ)	24.19	3.91	(.82)					
2) Behavioral inhibition (BISQ)	20.39	4.43	.04	(.86)				
3) Numeracy (BNT)	2.12	1.55	−.20*	−.10	(.80)			
4) Cognitive reflection (CRT)	3.43	2.00	−.24**	−.08	.79***	(.78)		
5) Discount rate gain	−2.69	1.10	.22*	.03	−.32***	−.40***	—	
6) Discount rate loss	−3.39	1.17	.03	.00	−.14	−.22*	.41***	—

Note. RRQ = Reward Responsiveness Questionnaire; BISQ = Behavioral Inhibition System Questionnaire; BNT = Berlin Numeracy Test; CRT = Cognitive Reflection Test. Discount rates expressed as log k of daily discount rates; reliability measured by Cronbach's α is presented in parentheses.
* $p < .05$. ** $p < .01$. *** $p < .001$.
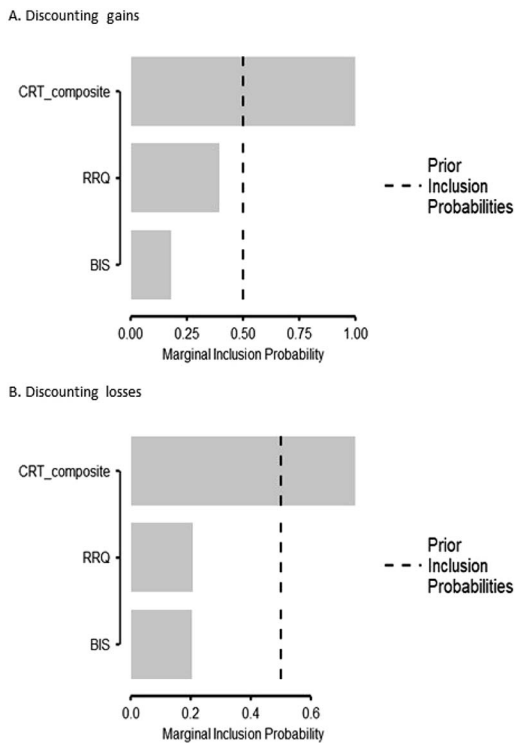
A. Discounting gains

B. Discounting losses

Figure 1. Marginal inclusion probabilities of all considered predictors for gains (Panel A) and for losses (Panel B). CRT = Cognitive Reflection Test; RRQ = Reward responsiveness; BISQ = Behavioral inhibition.

ginal inclusion probabilities for each of the standardized predictors.

For gains, the model best supported by the evidence included only the CRT ($\beta = -.400$), adjusted $R^2 = .153$, $F(1, 123) = 23.26$, $p < .001$, $B_M = 7.07$. However, the model assuming two main effects of cognitive reflection and the RRQ ($B_{01} = 1.58$) was almost as strongly supported by the data.[4]

For losses, the model best supported by the data is the one including only the CRT ($\beta = -.215$), adjusted $R^2 = .038$, $F(1, 123) = 5.90$, $p = .017$, $B_M = 6.08$. However, the null model ($B_{01} = 2.67$) cannot be validly refuted.

Finally, we tested for the possibility that the trait that overlaps in CRT and BNT (this may be numeracy that contributes to the CRT, but it may also be cognitive reflection that could contribute to the BNT, or some other third factor) actually predicts discount rates. To this end, we created a combined score of both scales and

correlated it with the discount rates (see Baron, Scott, Fincher, & Metz, 2015 for similar procedure). The reliability of the combined scales was higher than of any of the scales separately ($\alpha = .896$), but the correlations were slightly smaller compared with correlations of the CRT only, with $r(122) = -.386$, $p < .001$ for gains, and $r(122) = -.190$, $p = .034$ for losses. This suggests that combining CRT and BNT scales is justified but does not improve the predictive power of models compared with models including only the CRT.

Discussion

In this experiment, we found that cognitive reflection was the only consistent predictor of discount rates. Reinforcement sensitivity may also be involved in the discounting of gains but not independently from cognitive reflection. Surprisingly, behavioral inhibition, which is claimed to guide behavior toward losses, was not correlated to the strength of loss discounting.

Although suggestive, our findings cannot decisively support the superiority of cognitive style over temperamental factors in predicting impatience, as there are some confounds and limitations. Specifically, the CRT items we used were strongly correlated with the numeracy; therefore we cannot attribute the effects of the CRT uniquely to cognitive reflection. In addition, the BIS/BAS scale we used, despite its popularity, is more suitable to investigate the original reinforcement sensitivity theory (Harnett, Loxton, & Jackson, 2013). The revised RST (Gray & McNaughton, 2003) decomposes the BIS scale into two factors: Fight, Flight, Freeze System (FFFS), which detects threat and punishment and elicits the subjective experience of fear, and revised Behavioural Inhibition System (r-BIS), which is concerned with conflict detection and resolution and elicits anxiety (Harnett et al., 2013). To test the revised theory, a more suitable tool would be the Reinforcement Sensitivity Theory Personality Questionnaire (RST-PQ; Corr & Cooper, 2016). Another potential problem with Experiment 1 is that we used a fixed order for the discounting tasks, and out of 195 participants who finished the first

---

[4] Where $B < 3$ is labelled as inconclusive, and $B < 10$ as moderate evidence (Dienes, 2014).

discounting task (i.e., gains), only 157 finished the second task (i.e., losses). Therefore, attrition may be a confounding factor.

To investigate this possibility, we decided to conduct a second experiment, in which we added the classical cognitive reflection test (Frederick, 2005), another version of the CRT (Thomson & Oppenheimer, 2016), and replaced the RRQ and BISQ with the RST-PQ. We also counterbalanced the order in which the discounting tasks were administered. Experiment 2 is therefore a conceptual replication of Experiment 1, in which we tested the same hypothesis but differently operationalized the predictors.

Experiment 2

This experiment was preregistered in Open Science Framework (https://osf.io/fze75/). Data for this experiment are available at https://osf.io/8k69y/.

Participants

A total of 245 participants were recruited via TurkPrime (Litman et al., 2017), and they participated in exchange for $1. We used the same recruitment and exclusionary criteria as in Experiment 1. Participants who completed Experiment 1 were prevented from participating in this experiment. A total of 28 participants were removed from analyses because their discount rates were either zero or negative, and further 92 were removed because they failed to respond to all questions in our survey (mostly to some of the RST-PQ items). Ultimately, we analyzed data from $n = 125$ participants (50 female, 72 male, three missing, $M_{age} = 35.0$, range: 20–68, $SD = 9.7$).

Materials

We used identical discounting and numeracy tests as in Experiment 1. We used newer tools to assess cognitive reflection and reinforcement sensitivity, which are described below.

The Reinforcement Sensitivity Theory of Personality Questionnaire. The RST-PQ is a questionnaire designed to measure the revised reinforcement sensitivity theory of personality (Corr & Cooper, 2016). Instead of the originally suggested two-factor structure, it contains a six-factor structure: two unitary defensive factors,

FFFS (related to fear) and the BIS (related to anxiety), and four BAS factors, Reward Interest, Goal-Drive Persistence, Reward Reactivity, and Impulsivity. The RST-PQ consists of 65 self-reported statements such as "I would run fast if I knew someone was following me late at night."

Cognitive Reflection Test. Because the two CRT tests we used in Experiment 1 overlapped with numeracy to an excessive extent, we decided to add the original CRT (Frederick, 2005) and one other parallel version that was designed to reduce the contribution of numeracy (Thomson & Oppenheimer, 2016). This resulted in a total of 13 test items; the same six items from Experiment 1 plus the three items from Frederick (2005) and the four items from Thomson and Oppenheimer (2016). When administered, all test items were combined, with the chronological order of the CRT tests determining the order in which the items were presented to the participants (items from older CRT tests were first). Answers to 12 of these 13 items were open ended, with one of the items from Toplak et al. (2014) being multiple choice. We calculated a composite score of cognitive reflectivity by summing the correct answers in all 13 items.

Procedure

Participants responded via an online survey to the RST-PQ, BNT, CRT, and ToAD. These materials were delivered in the fixed sequential order mentioned earlier; however, this time, delay discounting tasks for gains and losses were counterbalanced. Participants were not able to return to previously answered sections of the survey.

Results

First, we found that the reliability of our measures (assessed with Cronbach's α) was similar to the reliability found in Experiment 1 and in the original articles. Also, the correlations between RST-PQ subscales (Table 2), despite being a bit stronger than originally reported, were always in the same directions. Finally, we again observed excessive correlations between two versions of the CRT (i.e., Primi et al., 2016; Toplak et al., 2014) and BNT. These have important implications for using alternative ver-

Table 2
Correlations Between Reinforcement Sensitivity, Cognitive Reflection, and Delay Discounting in Experiment 2

Dependent Variable	M	SD	1	2	3	4	5	6	7	8	9	10
1) Fight–flight–freeze system	23.37	6.74	(.84)									
2) Behavioral inhibition	48.45	15.83	.50***	(.95)								
3) Reward interest	17.67	4.64	.16	−.02	(.84)							
4) Goal-drive persistence	20.50	4.712	.17	−.19*	.53***	(.87)						
5) Reward reactivity	25.50	5.40	.38***	.23*	.59***	.51***	(.81)					
6) Impulsivity	15.96	4.65	.34***	.33***	.39***	−.14	.37***	(.79)				
7) Numeracy	1.94	1.45	−.33**	−.08	−.19*	−.05	−.06	−.21*	(.73)			
8) Cognitive reflection	8.14	3.75	−.24**	−.12	−.19*	.04	−.07	−.23*	.74***	(.87)		
9) Discount rate loss	−3.36	1.07	.09	.04	.10	.06	.10	.06	−.10	−.18*	—	
10) Discount rate gain	−2.59	1.31	.06	.06	.03	.04	.08	.03	−.16	−.18*	.37**	—

Note. Discount rates expressed as log k of daily discount rates; reliability measured by Cronbach's α is presented in parentheses.
* $p < .05$. ** $p < .01$. *** $p < .001$.

sions of the CRT. To conclude, despite a substantial number of individuals who failed to respond to all questions asked, the collected data appear valid and reliable.

Confirmatory Analysis

Our preregistered analysis required us to regress discounting rates on all predictors. Yet, because of the excessive BNT by CRT correlation, our analysis was limited to the main effects of all the six subscales of the RST-PQ, and the CRT, but did not include the BNT.

Consistent with Experiment 1, the best model to predict discounting of gains included only the cognitive reflection test ($\beta = -.180$), adjusted $R^2 = .025$, $F(1, 124) = 4.12$, $p = .045$, $B_M = 16.77$. However, the null model was almost equally supported by the data ($B_{01} = 1.23$).

For losses, the best model included only the CRT ($\beta = -.181$), adjusted $R^2 = .025$, $F(1, 124) = 4.18$, $p = .043$, $B_M = 12.17$. However, the null model was almost equally well supported by the evidence ($B_{01} = 1.03$) and only weakly better supported by the evidence compared with models including effects of cognitive reflection jointly with reward reactivity (RR; $B_{01} = 2.32$), or with reward interest (RI; $B_{01} = 2.70$), or with goal-drive persistence (GDP; $B_{01} = 2.76$), or with only the main effect of RI ($B_{01} = 2.79$) or RR ($B_{01} = 2.90$). Figure 2 presents marginal inclusion probabilities for gains (Panel A) and for losses (Panel B).

Exploratory Analysis

Because of the excessive BNT by CRT correlations, our confirmatory analysis did not include numeracy. Here, we reanalyzed data by independently using versions of the Frederick, and Thomson and Oppenheimer scales, as these were less strongly correlated to the BNT when compared with the composite score. These predictors were accompanied by the BNT. We also tested for composite score of BNT and CRTs.

Frederick CRT. For gains, the best model included only the CRT ($B_M = 21.29$), but it had almost equal evidential support as the null model ($B_{01} = 1.55$), or as the model with only the BNT ($B_{01} = 1.78$). For losses, the null model was most strongly supported by the data ($B_M = 24.44$), with anecdotal evidence supporting it over models including only the CRT ($B_{01} = 1.33$), only the RI ($B_{01} = 2.71$), and only the RR ($B_{01} = 2.82$).

Thomson CRT. For gains the null model was most strongly supported by the data ($B_M = 20.55$), with anecdotal evidence supporting it over model including only the BNT ($B_{01} = 1.14$), or over the model including only the CRT ($B_{01} = 2.13$). For losses, the best model included only the CRT ($B_M = 12.04$), with anecdotal evidence supporting it over null model ($B_{01} = 1.12$), and over models including CRT and RR ($B_{01} = 2.31$), CRT and RI ($B_{01} = 2.64$), CRT and GDP ($B_{01} = 2.74$), and CRT and FFFS ($B_{01} = 2.88$).
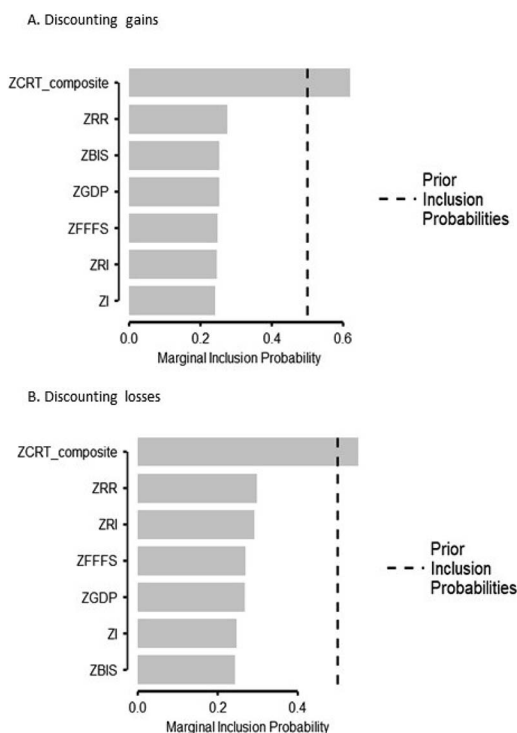
A. Discounting gains

B. Discounting losses

Figure 2. Marginal inclusion probabilities of all considered predictors for gains (Panel A) and for losses (Panel B). ZRR = Reward Reactivity; ZRI = Reward interest; ZFFFS = Fight–flight–freeze system; ZGDP = Goal-drive persistence; ZI = Impulsivity; ZBIS = Behavioral inhibition System. Z in each acronym denotes the scores were standardized.

Combined scores of BNT and CRT. For gains, the best model included only the combined score ($B_M = 17.31$), almost equal evidential support over the null model ($B_{01} = 1.40$). For losses the null model was most strongly supported by the data ($B_M = 12.77$), with anecdotal evidence supporting it over models including only one of the following: the combined score ($B_{01} = 1.27$), BAS ($B_{01} = 2.43$) RI ($B_{01} = 2.71$), or the RR ($B_{01} = 2.82$).

Discussion

With this experiment, we confirmed that cognitive reflection is the only consistent predictor of delay discounting, from the considered predictors. However, in this research, we found substantially weaker evidence for this claim compared with Experiment 1, with *p* values for discounting of gains and losses just below the

.05 threshold and lower bounds of the effect sizes' confidence intervals just exceeding zero. Finally, by using two versions of the CRT that were not as strongly correlated to the BNT, we found that cognitive reflection is a better than numeracy at predicting delay discounting behavior.

General Discussion

The two reported experiments suggest that cognitive reflection, but not reinforcement sensitivity or numeracy, is the only consistent predictor of discount rates in gains and losses. The reason why cognitively reflective individuals tend to be less impatient in their intertemporal choice can be attributed to the way people process information, with more reflective individuals being better at overriding their initial intuitions. However, because numeracy's predictive power was comparable with the CRT, and their combined scores predicted discount rates just as well as these scales independently, it may be the case that numeracy is what actually predicts discount rates. The alternative explanation, suggesting that more patient individuals are simply less affected by gains and losses had little to no support in this research. Our findings are robust because they are consistent for both gains and losses and across four different measures of cognitive reflection. Our findings are also valuable as part of them comes from a preregistered protocol, where we explicitly stated our predictions and described the methods we planned to use to test them.

Establishing that cognitive reflection, but not reinforcement sensitivity, is predictive of intertemporal choice provides us with a better understanding of why people are impatient. In gains, people tend to accelerate their payoffs even if it means receiving less. Classically, peoples' impulsive decisions (such as eating junk food) are attributed to the distorted attractiveness of rewards. In other words, it was assumed that people tend to overeat because consuming food is extremely pleasurable (Lowe & Butryn, 2007). Yet, our findings show that subjective attractiveness of a gain does not predict discount rates for gains, most likely because it fails at accurately capturing the actual subjective value of a gain. According to our data, a person's ultimate intertemporal choice depends on whether they can override the impulsive urge to

receive gains immediately (which depends on cognitive reflection).

In losses, people with strong discount rates tend to postpone losses even when this means losing much more. For example, people might postpone visiting a dentist to fill a small cavity, which may eventually lead to a more serious issue that requires them to undergo a root canal procedure. Like discounting of gains, our data suggests that discounting of losses is not explained by the subjective sensitivity to losses. Instead, those with greater cognitive reflection will likely be able to override their impulses and make more reasoned decisions, expressed here as weaker discounting.

It is sometimes claimed that discount rates in gains and losses depend on different cognitive processes, with some of the studies reporting that cognitive reflection does not affect the discounting of losses (Frederick, 2005; Hardisty & Weber, 2009). In this research, we observed cognitive reflection to be predictive of discounting in both gains and losses, but its impact is possibly stronger in gains (e.g., in Experiment 1, the effect of cognitive reflection was two times stronger for predicting discount rates in gains than in losses, but we observed no such difference in Experiment 2). If this is true and if cognitive reflection is more strongly tied to discounting of gains, then cognitive reflection's effect on the discounting of losses requires more power and are more difficult to capture experimentally. This would explain why previous studies failed to find cognitive reflection effects in losses: The same sample size has much greater chances to detect stronger effects (Simonsohn, 2015). Despite the potential difference in the strength of CRT's impact on gains and losses, our results provide evidence that high discount rates in both can be effectively countered by cognitive reflection.

A limitation of the current research is that the incremental validity approach adapted in this research, although widely used, suffers from some problems. Specifically, because of the imperfect reliability of the measurements, the chances of false discoveries are higher than previously assumed (Westfall & Yarkoni, 2016). Therefore, it may be the case that our predictors share some common variance, and, in turn, this inflates the impact of cognitive reflection in multiple regression. However, independent correlations (Tables 1 and 2) show that

only the CRT consistently correlates with discount rates, validating our conclusions.

Another limitation of our research is the use of a particular method for eliciting discount rates: hypothetical monetary choices. Though discount rates elicited using similar methods are validated as a predictor of several psychological issues such as gambling or obesity, one might want to test whether cognitive reflection indeed predicts discounting of nonhypothetical or nonfinancial outcomes. There exists some evidence that delay discounting might be domain specific (Holt, Glodowski, Smits-Seemann, & Tiry, 2016; Holt, Newquist, Smits, & Tiry, 2014; Tsukayama & Duckworth, 2010). Domain specificity may be also understood as an incarnation of magnitude effect, in which payoffs from some domains (e.g., sex) are more valuable than for other domains (e.g., food) and therefore are discounted less strongly (Sawicki, Markiewicz, & Białek, 2019). Next, reinforcement sensitivity could be affecting discount rates only when considered payoffs are real rather than hypothetical. Hence the question of whether or not our findings can be extended to real-life hedonic choices, such as junk food or casual sex, requires further empirical investigation. Until that evidence is provided, we are limiting our conclusions to just hypothetical monetary decisions.

Finally, there is a broader issue with online-recruited, paid subject samples. Some participants may be self-selected because of their particular traits (e.g., they think they perform well in cognitive tasks). Some other participants may be especially vulnerable to the complexity of the experimental task and end up being rejected (e.g., participants who lack cognitive reflection may also fail at comprehension checks or to compete all tasks assigned). Note that if this was true, our experiment would underestimate the importance of cognitive reflection because data from those scoring the lowest on the CRT are most likely to be lost at the data reduction stage. Future research might avoid this issue by running lab-based experiments in which participants are under the supervision of research assistants. This helps participants to stabilize their top-down control and, in turn, can reduce performance fluctuations (Steinborn & Huestegge, 2019). Experimental design could also mobilize participants by instructing them

to "try hard" when running critical trials of the experiment (Steinborn, Langner, & Huestegge, 2017), that is, the delay discounting procedure. This should reduce drop-out rates in low-reflective participants and ensure higher-quality samples. Moreover, one could also control for subjective energy and stress level before and after the study, as it is increasingly evident it critically affects behavior and performance in both questionnaires and performance tests. To this end, the Dundee Stress State Questionnaire (Langner, Steinborn, Chatterjee, Sturm, & Willmes, 2010; Matthews et al., 2002) could be implemented in future research.

Our results exposed a strong relationship between cognitive reflection and numeracy. This relationship varied in strength for different versions of the CRT, suggesting that using different CRT tests might lead to different findings, as some of them are almost unequivocal tests of numeracy (Sirota et al., 2018). Researchers interested in cognitive reflection should consider careful selection of the items they use, as even using composite tests by selecting particular items results in high correlations with numeracy. Researchers might consider continuing to use the original CRT (Frederick, 2005), as familiarity with the test does not necessarily increase the raw score of participants (Meyer, Zhou, & Frederick, 2018), nor have any negative effect on the test validity (Białek & Pennycook, 2018).

To sum up, we provide evidence that strong discounting is a result of how people combine various goals rather than the strength of a particular goal. In other words, it is not so critical to know how strong a temptation is but rather to what extent one is willing to override it with other thoughts and goals. What drives opting for an immediate $100 gain is not how much one needs this money now, but to what extent one is willing to also consider how much more can be obtained in the future. Therefore, the most promising way of changing behavior is likely focusing on changing the cognitive style underlying decision-making. We call on researchers to actively search for such interventions, as there are currently none available, because they would likely greatly benefit the quality of peoples' everyday lives.

References

Abdellaoui, M., Attema, A. E., & Bleichrodt, H. (2010). Intertemporal tradeoffs for gains and losses: An experimental measurement of discounted utility. *Economic Journal, 120,* 845–866. http://dx.doi.org/10.1111/j.1468-0297.2009.02308.x

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. http://dx.doi.org/10.1007/s11166-011-9125-1

Baron, J., Scott, S., Fincher, K., & Metz, S. E. (2015). Why does the Cognitive Reflection Test (sometimes) predict utilitarian moral judgment (and other things)? *Journal of Applied Research in Memory and Cognition, 4,* 265–284. http://dx.doi.org/10.1016/j.jarmac.2014.09.003

Benhabib, J., Bisin, A., & Schotter, A. (2010). Present-bias, quasi-hyperbolic discounting, and fixed costs. *Games and Economic Behavior, 69,* 205–223. http://dx.doi.org/10.1016/j.geb.2009.11.003

Benningfield, M. M., Blackford, J. U., Ellsworth, M. E., Samanez-Larkin, G. R., Martin, P. R., Cowan, R. L., & Zald, D. H. (2014). Caudate responses to reward anticipation associated with delay discounting behavior in healthy youth. *Developmental Cognitive Neuroscience, 7,* 43–52. http://dx.doi.org/10.1016/j.dcn.2013.10.009

Bettman, J. R., Luce, M. F., & Payne, J. W. (1998). Constructive consumer choice processes. *Journal of Consumer Research, 25,* 187–217. http://dx.doi.org/10.1086/209535

Białek, M., & Pennycook, G. (2018). The Cognitive Reflection Test is robust to multiple exposures. *Behavior Research Methods, 50,* 1953–1959. http://dx.doi.org/10.3758/s13428-017-0963-x

Białek, M., & Sawicki, P. (2018). Cognitive reflection effects on time discounting. *Journal of Individual Differences, 39,* 99–106. http://dx.doi.org/10.1027/1614-0001/a000254

Bickel, W. K., Odum, A. L., & Madden, G. J. (1999). Impulsivity and cigarette smoking: Delay discounting in current, never, and ex-smokers. *Psychopharmacology, 146,* 447–454. http://dx.doi.org/10.1007/PL00005490

Büchel, C., Peters, J., Banaschewski, T., Bokde, A. L., Bromberg, U., Conrod, P. J., . . . Knutson, B. (2017). Blunted ventral striatal responses to anticipated rewards foreshadow problematic drug use in novelty-seeking adolescents. *Nature Communications, 8.* http://dx.doi.org/10.1038/ncomms14140

Campitelli, G., & Gerrans, P. (2014). Does the cognitive reflection test measure cognitive reflection? A mathematical modeling approach. *Memory and Cognition, 42,* 434–447. http://dx.doi.org/10.3758/s13421-013-0367-9

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS Scales. *Journal of Personality and Social Psychology, 67,* 319–333. http://dx.doi.org/10.1037/0022-3514.67.2.319

Chapman, G. B. (1996). Temporal discounting and utility for health and money. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 22,* 771–791. http://dx.doi.org/10.1037//0278-7393.22.3.771

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

Corr, P. J. (2004). Reinforcement sensitivity theory and personality. *Neuroscience and Biobehavioral Reviews, 28,* 317–332. http://dx.doi.org/10.1016/j.neubiorev.2004.01.005

Corr, P. J., & Cooper, A. J. (2016). The Reinforcement Sensitivity Theory of Personality Questionnaire (RST-PQ): Development and validation. *Psychological Assessment, 28,* 1427–1440. http://dx.doi.org/10.1037/pas0000273

Corvi, A. P., Juergensen, J., Weaver, J. S., & Demaree, H. A. (2012). Subjective time perception and behavioral activation system strength predict delay of gratification ability. *Motivation and Emotion, 36,* 483–490. http://dx.doi.org/10.1007/s11031-011-9275-0

Daugherty, J. R., & Brase, G. L. (2010). Taking time to be healthy: Predicting health behaviors with delay discounting and time perspective. *Personality and Individual Differences, 48,* 202–207. http://dx.doi.org/10.1016/j.paid.2009.10.007

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology, 5,* 781. http://dx.doi.org/10.3389/fpsyg.2014.00781

Estle, S. J., Green, L., Myerson, J., & Holt, D. D. (2006). Differential effects of amount on temporal and probability discounting of gains and losses. *Memory and Cognition, 34,* 914–928. http://dx.doi.org/10.3758/BF03193437

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

Gray, J. A. (1982). Précis of the neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system. *Behavioral and Brain Sciences, 5,* 469–484. http://dx.doi.org/10.1017/S0140525X00013066

Gray, J. A., & McNaughton, N. (2003). *The neuropsychology of anxiety: An enquiry into the function of the septo-hippocampal system.* New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780198522713.001.0001

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

Hansen, A., Turpyn, C. C., Mauro, K., Thompson, J. C., & Chaplin, T. M. (2019). Adolescent brain response to reward is associated with a bias toward immediate reward. *Developmental Neuropsychology, 44,* 417–428. http://dx.doi.org/10.1080/87565641.2019.1636798

Hardisty, D. J., & Weber, E. U. (2009). Discounting future green: Money versus the environment. *Journal of Experimental Psychology: General, 138,* 329–340. http://dx.doi.org/10.1037/a0016433

Harnett, P. H., Loxton, N. J., & Jackson, C. J. (2013). Revised reinforcement sensitivity theory: Implications for psychopathology and psychological health. *Personality and Individual Differences, 54,* 432–437. http://dx.doi.org/10.1016/j.paid.2012.10.019

Held, L., & Ott, M. (2018). On p-values and Bayes factors. *Annual Review of Statistics and Its Application, 5,* 393–419. http://dx.doi.org/10.1146/annurev-statistics-031017-100307

Hirsh, J. B., Morisano, D., & Peterson, J. B. (2008). Delay discounting: Interactions between personality and cognitive ability. *Journal of Research in Personality, 42,* 1646–1650. http://dx.doi.org/10.1016/j.jrp.2008.07.005

Holt, D. D., Glodowski, K., Smits-Seemann, R. R., & Tiry, A. M. (2016). The domain effect in delay discounting: The roles of fungibility and perishability. *Behavioural Processes, 131,* 47–52. http://dx.doi.org/10.1016/j.beproc.2016.08.006

Holt, D. D., Newquist, M. H., Smits, R. R., & Tiry, A. M. (2014). Discounting of food, sex, and money. *Psychonomic Bulletin and Review, 21,* 794–802. http://dx.doi.org/10.3758/s13423-013-0557-2

Jarmolowicz, D. P., Cherry, J. B. C., Reed, D. D., Bruce, J. M., Crespi, J. M., Lusk, J. L., & Bruce, A. S. (2014). Robust relation between temporal discounting rates and body mass. *Appetite, 78,* 63–67. http://dx.doi.org/10.1016/j.appet.2014.02.013

Kidd, C., Palmeri, H., & Aslin, R. N. (2013). Rational snacking: Young children's decision-making on the marshmallow task is moderated by beliefs about environmental reliability. *Cognition, 126,* 109–114. http://dx.doi.org/10.1016/j.cognition.2012.08.004

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. http://dx.doi.org/10.1016/j.actpsy.2009.10.001

Liberali, J. M., Reyna, V. F., Furlan, S., Stein, L. M., & Pardo, S. T. (2012). Individual differences in numeracy and cognitive reflection, with implica-

tions for biases and fallacies in probability judgment. *Journal of Behavioral Decision Making, 25,* 361–381. http://dx.doi.org/10.1002/bdm.752

Litman, L., Robinson, J., & Abberbock, T. (2017). TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods, 49,* 433–442. http://dx.doi.org/10.3758/s13428-016-0727-z

Loewenstein, G. (1987). Anticipation and the valuation of delayed consumption. *Economic Journal, 97,* 666–684. http://dx.doi.org/10.2307/2232929

Loewenstein, G. F. (1988). Frames of mind in intertemporal choice. *Management Science, 34,* 200–214. http://dx.doi.org/10.1287/mnsc.34.2.200

Lowe, M. R., & Butryn, M. L. (2007). Hedonic hunger: A new dimension of appetite? *Physiology and Behavior, 91,* 432–439. http://dx.doi.org/10.1016/j.physbeh.2007.04.006

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. http://dx.doi.org/10.1007/s00213-011-2229-0

Matthews, G., Campbell, S. E., Falconer, S., Joyner, L. A., Huggins, J., Gilliland, K., . . . Warm, J. S. (2002). Fundamental dimensions of subjective state in performance settings: Task engagement, distress, and worry. *Emotion, 2,* 315–340. http://dx.doi.org/10.1037/1528-3542.2.4.315

Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *Quantitative analyses of behavior: Vol. 5. The effect of delay and of intervening events on reinforcement value* (pp. 55–73). Hillsdale, NJ: Erlbaum.

McClure, S. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science, 306,* 503–507. http://dx.doi.org/10.1126/science.1100907

Mellick, W., Tolliver, B. K., Brenner, H., & Prisciandaro, J. J. (2019). Delay discounting and reward sensitivity in a 2 × 2 study of bipolar disorder and alcohol dependence. *Addiction, 114,* 1369–1378. http://dx.doi.org/10.1111/add.14625

Meyer, A., Zhou, E., & Frederick, S. (2018). The non-effects of repeated exposure to the cognitive reflection test. *Judgment and Decision Making, 13,* 246–259.

Mishra, S., & Lalumiere, M. L. (2017). Associations between delay discounting and risk-related behaviors, traits, attitudes, and outcomes. *Journal of Behavioral Decision Making, 30,* 769–781. http://dx.doi.org/10.1002/bdm.2000

Mitchell, S. H., & Wilson, V. B. (2010). The subjective value of delayed and probabilistic outcomes: Outcome size matters for gains but not for losses. *Behavioural Processes, 83,* 36–40. http://dx.doi.org/10.1016/j.beproc.2009.09.003

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., . . . Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences of the United States of America, 108,* 2693–2698. http://dx.doi.org/10.1073/pnas.1010076108

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

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. http://dx.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. http://dx.doi.org/10.3389/fpsyg.2016.00009

Perry, J. L., Larson, E. B., German, J. P., Madden, G. J., & Carroll, M. E. (2005). Impulsivity (delay discounting) as a predictor of acquisition of IV cocaine self-administration in female rats. *Psychopharmacology, 178,* 193–201. http://dx.doi.org/10.1007/s00213-004-1994-4

Peters, J., & Büchel, C. (2011). The neural mechanisms of inter-temporal decision-making: Understanding variability. *Trends in Cognitive Sciences, 15,* 227–239. http://dx.doi.org/10.1016/j.tics.2011.03.002

Primi, C., Morsanyi, K., Chiesi, F., Donati, M. A., & Hamilton, J. (2016). The development and testing of a new version of the cognitive reflection test applying item response theory (IRT). *Journal of Behavioral Decision Making, 29,* 453–469. http://dx.doi.org/10.1002/bdm.1883

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

Sawicki, P., & Białek, M. (2017). Overlapping defaults. The case of intertemporal choices. *Polish Psychological Bulletin, 48,* 440–444. http://dx.doi.org/10.1515/ppb-2017-0050

Sawicki, P., Markiewicz, Ł., & Białek, M. (2019). Magnitude effect contributes to the domain specificity in delay discounting. *Journal of Behavioral Decision Making.* Retrieved form https://www.researchgate.net/profile/Michal_Bialek2/publication/335603324_Magnitude_effect_contributes_to_the_domain_specificity_in_delay_discounting/links/5d6fb5bc4585151ee49b9650/Magnitude-effect-contributes-to-the-domain-specificity-in-delay-discounting.pdf

Shamosh, N. A., Deyoung, C. G., Green, A. E., Reis, D. L., Johnson, M. R., Conway, A. R. A., . . . Gray, J. R. (2008). Individual differences in delay discounting: Relation to intelligence, working memory, and anterior prefrontal cortex. *Psychological Science, 19,* 904–911. http://dx.doi.org/10.1111/j .1467-9280.2008.02175.x

Shamosh, N. A., & Gray, J. R. (2008). Delay discounting and intelligence: A meta-analysis. *Intelligence, 36,* 289–305. http://dx.doi.org/10.1016/j .intell.2007.09.004

Siegelman, M. E. (2016). *Construction of an effort discounting task (EDT). Honors Theses - All. 1584*. Retrieved from https://wesscholar.wesleyan .edu/etd_hon_theses/1584

Simonsohn, U. (2015). Small telescopes: Detectability and the evaluation of replication results. *Psychological Science, 26,* 559–569. http://dx.doi.org/ 10.1177/0956797614567341

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

Sirota, M., Kostovičová, L., Juanchich, M., Dewberry, C., & Marshall, A. C. (2018). Measuring cognitive reflection without maths: Developing and validating the verbal cognitive reflection test. *PsyArxiv*. Advanced online publication. http://dx .doi.org/10.31234/osf.io/pfe79

Stanovich, K. E. (2018). Miserliness in human cognition: The interaction of detection, override and mindware. *Thinking and Reasoning, 24,* 423–444. http://dx.doi.org/10.1080/13546783.2018.1459314

Stanovich, K. E., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology, 94,* 672–695. http://dx.doi.org/10.1037/0022-3514 .94.4.672

Steinborn, M. B., & Huestegge, L. (2019). Socially alerted cognition evoked by a confederate's mere presence: Analysis of reaction-time distributions and delta plots. *Psychological Research*. Advanced online publication. http://dx.doi.org/10.1007/s00426- 019-01143-z

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. http://dx.doi.org/10.1007/s00426- 016-0810-1

Story, G. W., Vlaev, I., Seymour, B., Winston, J. S., Darzi, A., & Dolan, R. J. (2013). Dread and the disvalue of future pain. *PLoS Computational Biology, 9,* e1003335. http://dx.doi.org/10.1371/journal .pcbi.1003335

Thaler, R. H. (1981). Some empirical evidence on dynamic inconsistency. *Economics Letters, 8,* 201–207. http://dx.doi.org/10.1016/0165-1765(81)90067-7

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

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the cognitive reflection test. *Thinking and Reasoning, 20,* 147–168. http://dx .doi.org/10.1080/13546783.2013.844729

Tsukayama, E., & Duckworth, A. L. (2010). Domain-specific temporal discounting and temptation. *Judgment and Decision Making, 5,* 72–82.

Van den Berg, I., Franken, I. H., & Muris, P. (2010). A new scale for measuring reward responsiveness. *Frontiers in Psychology, 1,* 239. http://dx.doi.org/ 10.3389/fpsyg.2010.00239

Van den Bergh, B., Dewitte, S., & Warlop, L. (2008). Bikinis instigate generalized impatience in intertemporal choice. *Journal of Consumer Research, 35,* 85–97. http://dx.doi.org/10.1086/525505

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. http://dx.doi.org/10.1111/j.1467- 9280.2007.01932.x

Westfall, J., & Yarkoni, T. (2016). Statistically controlling for confounding constructs is harder than you think. *PLoS ONE, 11,* e0152719. http://dx.doi .org/10.1371/journal.pone.0152719

Yoon, H., & Chapman, G. B. (2016). A closer look at the yardstick: A new discount rate measure with precision and range. *Journal of Behavioral Decision Making, 29,* 470–480. http://dx.doi.org/10.1002/bdm .1890