Title: A Computer Adaptive Measure of Delay Discounting

Authors: Vaishali Mahalingam[1], Michael Palkovics[2], Michal Kosinski[3], Iva Cek[1] & David Stillwell[4]

Affiliations:

[1] Department of Psychology, University of Cambridge, UK.

[2] Department of Philosophy and Education, University of Vienna, Austria.

[3] Graduate Business School, Stanford University, US;

[4] Cambridge Judge Business School, University of Cambridge, UK.

Corresponding Author: Vaishali Mahalingam, The Psychometrics Centre, Department of Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, UK.

Email: vm298@cam.ac.uk

Declarations of competing interest: None

Acknowledgements:

V.M. was supported by the Nehru Trust for Cambridge University. I.C. was supported by the Cambridge Overseas Trust. We thank Dr. Igor Gomes Menezes for his critical reading of the manuscript.

Words: 7638

Figures: 4

Tables: 9

References: 65

Abstract

Delay discounting has been linked to important behavioral, health and social outcomes, including academic achievement, social functioning and substance use, but thoroughly measuring delay discounting is tedious and time consuming. We develop and consistently validate an efficient and psychometrically sound computer adaptive measure of discounting. First, we develop a binary search-type algorithm to measure discounting using a large international dataset of 4,190 participants. Using six independent samples (N=1550), we then present evidence of concurrent validity with two standard measures of discounting and a measure of discounting real rewards, convergent validity with addictive behavior, impulsivity, personality, survival probability; and divergent validity with time perspective, life satisfaction, and age and gender. The new measure is considerably shorter than standard questionnaires, includes a range of time delays, can be applied to multiple reward magnitudes, shows excellent concurrent, convergent, divergent, and discriminant validity – by showing more sensitivity to effects of smoking behavior on discounting.

Keywords: delay discounting, computer adaptive testing, item response theory, hierarchical linear modelling/multilevel modelling, addiction, social network data.

Introduction

Would you rather have $650 now or $1000 in a month? $850 now or $1000 in a year?

Most people are impatient and so they subjectively devalue rewards as the delay before receiving them increases. Preferring smaller rewards has been associated with various negative outcomes: obesity (Weller, Cook, Avsar, & Cox, 2008), smoking (Bickel, Odum, & Madden, 1999; Johnson & Bickel, 2002; Krishnan-Sarin et al., 2007; Brady Reynolds et al., 2007) and drug use (Johnson & Bickel, 2002; Kirby & Petry, 2004; Madden et al., 2004). In contrast, more patient delay discounting has been linked to positive academic, health and social outcomes (Hirsh, Morisano, & Peterson, 2008; Kirby, Winston, & Santiesteban, 2005).

Established measures of delay discounting tend to be tedious and time consuming, requiring participants to respond to a large number of dichotomous items with a range of time delays and delayed amounts. Using a large international sample (N = 4,190), we develop a binary search-like algorithm to measure delay discounting and present the results of a simulation study comparing the new algorithm to item response theory-based computer adaptive testing and a standard measure. We then present evidence of concurrent, convergent, divergent and discriminant validity (N = 1550) for the newly developed computer adaptive measure of delay discounting.

Measurement of delay discounting

Standard versus adaptive measures

Standard measures of delay discounting measure the point of inflexion, where an individual switches from preferring a larger future reward to a smaller immediate reward, indicating an estimate of the individual's subjective value of the delayed reward relative to current monetary values. For example, an individual prefers $1000 in a year over $800 now, but also prefers $850 now over $1000 in a year. Here, the immediate subjective monetary value of $1000 in a year is between $800 and $850.

Traditionally, experimenters offered participants a series of binary choices in which delay lengths were predetermined (Green, Fry, & Myerson, 1994). This method becomes increasingly onerous as more time points are measured and at increasing accuracy. Rachlin, Raineri, & Cross (1991) measured seven time points (1 month - 50 years) and used 30 different immediate amounts ($1,000 - $1), amounting to 210 items. However, the delayed amount was always $1000, and so the test would be even longer if different delayed amounts were measured.

Experimental evidence shows that using different time delays and delayed amounts alters discounting rates (Lane, Cherek, Pietras, & Tcheremissine, 2003; Stillwell & Tunney, 2012); thus, measurement of a single indifference point for an arbitrary delay and amount is inadequate. A *magnitude effect* is found, whereby smaller rewards are discounted more steeply than larger ones (Green, Fristoe, & Myerson, 1994; Green, Fry, et al., 1994; Kirby, 1997; Mahalingam, Stillwell, Kosinski, Rust, & Kogan, 2014; Raineri & Rachlin, 1993). There is also an effect of delay time (Mahalingam et al., 2014; Stillwell & Tunney, 2012) and so studies use a range of time delays, from a few hours to 25 years, and average the estimated discounting parameter at each delay.

In the past, researchers have attempted to counter the length of traditional delay discounting measures. Reimers et al. (2009) used a single item (£45 in three days versus £70 in three months) to measure discounting in a very large sample of 42,863 U.K. residents. Although this method is extremely quick and best suited for large samples, it is not possible to measure the point of inflexion without an upper and lower bound on the preference for the delayed reward. If the participant chooses the earlier reward, they subjectively value £70 in three months less than £45 in three days, but it is unclear whether the value is nearer £40, £20 or indeed £1. Similarly, the Monetary Choice Questionnaire (Kirby & Marakovic, 1996) also does not measure the point of inflexion. MCQ respondents make 21 binary choices between rewards at predefined delays (10-75 days) and amounts ($15-$85). Since delays and amounts are not tested multiple times, participants can only be ranked against one another within each questionnaire. Additionally, it is not simple to extrapolate beyond the specific delays and amounts tested, since they are scored as independent items by summing up earlier versus later responses rather than by calculating a delay discounting parameter.

In the Delay Discounting Index (Griskevicius, Tybur, Delton, & Robertson, 2011) seven dichotomous items measuring $100 tomorrow versus $100-$170 90 days from now are combined with two Likert items on a 9-point scale measuring the strength of preference between a smaller amount tomorrow versus a larger amount in either 3 or 12 months. While the seven dichotomous items can be used to establish a point of inflexion for the value of $100 in 90 days in present day terms, it is unclear how this can be combined with the Likert items, and so authors typically do not calculate the point of inflexion.

Attempts have been made to develop a computerized task for delay discounting. Johnson (2012) developed an operant choice procedure to *behaviorally* measure discounting. The task was designed to determine an individual's discounting function within 20 minutes, while establishing the indifference point on each time delay. The task consisted of 5 blocks of 4 trials, and accuracy of estimating indifference points increased with each trial. As participants engaged in the task in *real time*, time delays and reward magnitudes were limited to small delays and rewards (< 80 seconds, < 40 cents).

Richards et al. (1999) developed a discounting task in which participants made choices between a smaller immediate reward versus $10 that was delayed for 0-365 days. The adjusting amount procedure used to establish the indifference point meant that the next step was randomly selected based on predefined rules. This method was the most computer adaptive strategy that also prevented users from predicting amounts. However, it was relatively item heavy (median number of items per participant = 74) and so, our aim of identifying an efficient measure is unsatisfied.

Mathematical models of delay discounting

Standard Economic theory uses an exponential discounting function, implying time consistent discounting. Exponential discounting was initially proposed by Samuelson (1937) in the context of a larger theoretical framework about the measurement of utility, rather than as a normative function. However, it was widely adopted by economists over the years.

Time consistent discounting suggests that rate of discounting remains constant across time and is not affected by the delay in receiving a reward. That is, individuals will discount outcomes available today compared to tomorrow (e.g. $100 today or $110 tomorrow) in the same way that they will discount outcomes available today compared to next year (e.g. $100 today or $110 in 1 year). Research shows that this is in fact untrue, and that the rate of discounting varies as a function of the delay and reward magnitude.

A hyperbolic function appears to best explain delay discounting in humans because it accounts for such time inconsistent discounting (Rachlin et al., 1991; Takahashi, Ikeda, & Hasegawa, 2007). For example, people are likely to prefer $1000 in 1 year and 1 day over $990 in 1 year, but will prefer $990 immediately rather than $1000 tomorrow; short delays have a relatively greater impact than longer delays. The hyperbolic delay discounting function describes this preference by accounting for the effect of the length of delay, and fits individuals' discounting data better than the exponential function (Johnson & Bickel, 2002; Kirby, 1997; Rachlin et al., 1991).

$$V = \frac{A}{(1 + kD)}$$

Equation 1 Hyperbolic function

Although the majority of researchers today conform to the assumptions of hyperbolic discounting, a number of other models have been used in the past – refer Doyle (2013). Alternatively, Area Under the Curve (AUC) is an atheoretical method that is used to summarise points of inflexion (Myerson, Green, & Warusawitharana, 2001; Odum, 2011) where larger AUC represents less discounting (i.e. less impatience for delayed outcomes), while less AUC represents greater discounting.

Hypothetical versus real rewards

Experimenters have used actual rewards versus hypothetical rewards in delay discounting tasks on the premise that actual behavior may vary compared to hypothetical behavior; and so, incentivizing the task with real rewards will result in a more accurate representation of actual decision making behavior. However, the bulk of research evidence directly comparing hypothetical and real rewards shows that there is no significant difference in delay discounting between the two methods (Johnson & Bickel, 2002; Lagorio & Madden, 2005; Lawyer, Schoepflin, Green, & Jenks, 2011; Locey, Jones, & Rachlin, 2011; Madden, Begotka, Raiff, & Kastern, 2003; Madden et al., 2004; Magen, Dweck, & Gross, 2008; Matusiewicz, Carter, Landes, & Yi, 2013).[1]

[1] A review by Kirby (1997) suggested that the differences between actual and hypothetical rewards are of magnitude rather than shape of the delay discounting function.

Study 1: Development of the Computerised Task

Computer adaptive testing (CAT) is increasingly being used for the development and scoring of modern psychometric tools (Magis & Mahalingam, 2015; Wainer, 2010). CAT works such that as long as the participant responds correctly to items that are being presented, she will be presented with harder items that match her ability; while if the participant responds incorrectly to an item, she will be presented with an easier item. For delay discounting, if a participant has indicated that her current subjective value of $1000 in 1 month is higher than $500, we can reduce measurement time by not asking if she values it at magnitudes below $500.

The core success of CAT is the ability to tailor them to individuals, even on massive scales (Chang, 2015) as each individual only receives questions pertaining to their ability level. It has been shown that for large populations of test-takers, adaptive tests provide more precise and equally valid scores (e.g. Johnson & Weiss, 1980; Kingsbury & Weiss, 1980; Thissen & Mislevy, 2000) while reducing the number of items required by 50% (e.g Weiss & Kingsbury, 1984; Mardberg & Carlstedt, 1998; Moreno & Segall, 1997). The ability of CAT to replace traditional methods of testing has been demonstrated in various applications, such as assessment of anxiety (e.g. Gibbons et al., 2014; Walter et al., 2007), depression (e.g. Fliege et al., 2005; Gardner et al., 2004; Smits, Cuijpers, & van Straten, 2011), drug susceptibility (Kirisci et al., 2012) and personality (traits) (e.g. Forbey & Ben-Porath, 2007; Hol, Vorst, & Mellenbergh, 2008; Rudick, Yam, & Simms, 2013).

The typical drawbacks of IRT-based CAT – not being suited for open-ended questions and the need to calibrate items (Chang, 2015) – do not apply in our case. Others, such as requiring a sufficient number of test-runs for item-calibration, have been successfully addressed (as demonstrated below). Likewise, we do not expect a mode effect (as described by Alderson, 2000) to affect our results, as our participants can be described as being computer-literate. For our purpose, the reduction in item-exposure and time required for test-taking are potentially beneficial. It can prevent test-takers from getting into the habit of blindly selecting responses because of initial questions seeming “obvious”. E.g. not taking $10 now as opposed to $1000 next month, not taking $50 now as opposed to $1000 next month, and so on.

Comparing Different Methods of Adaptive Testing using Simulations

A simulation study was conducted to compare a binary search-type algorithm with item response theory-based computer adaptive testing and a standard measure of delay discounting. The goal was to accurately measure delay discounting while limiting the number of questions.

A subset of *N* = 4,190 from the myPersonality database of *N* = 19,202 participants (Stillwell & Kosinski, 2011) was chosen by excluding incomplete protocols, only including participants who took the questionnaire items in a randomized order and excluding inconsistent and extreme (+/- 3 SD) responses. Points of indifference were calculated for each participant and timeframe. Next, the percentiles for each individual timeframe were also calculated based on the data. This allowed us to later compare the actual percentile of a participant’s score with the score resulting out of the simulated process to establish which yielded more accurate results.

Binary Search-Like Algorithm

In the process of developing a suitable computer adaptive task, a number of possible methods were drafted and tested, at early stages. In one promising case, the test was made computer adaptive using a binary search-type algorithm to establish the switching point between receiving a reward now or at a specific point in the future. A binary search is an efficient way of searching a large array of sorted elements (Cormen, Leiserson, & Rivest, 1990; Horvath, 2012). At each step, the range of values is split in half by the middle percentile; in other words, the first question administered is at the 50th percentile and the second question administered is at the 75th or 25th percentile depending upon the participant's response to the first question.

Since indifference points are correlated between delays, the binary search can be improved when measuring many time delays by starting at the final indifference point from the previous trial rather than the 50th percentile. In this case, we can be fairly confident that the participant's indifference point will be close to the indifference point of the previous trial. Therefore, we used standard deviation units (16 percentiles) to pick the next percentile rather than the binary search. Once the participant indicates a preference reversal, we apply the binary search from then on.

For example, take an individual who was at the 25th percentile in the first time delay and whose hidden preference is at the 50th percentile for the next delay. He is first presented an item at the 25th percentile for the next time delay. At this point, depending on whether he opts for the immediate or delayed reward, the next item will be -/+ 16 percentiles away. As he prefers the delayed reward, he will be presented with the item at the 41st percentile. Since his real preference is around the 50th percentile, he then opts for the delayed amount again so that the 57th item is presented to him. As this is higher than his actual preference, he indicates a preference reversal by selecting the immediate amount. The binary search algorithm will then be used to select the following items. The next item to be presented would be between the 57th and 41st percentiles.

Item Response Theory (IRT)

IRT models allows for test characteristics, such as item difficulty and discrimination, and participant characteristics, such as ability and inattention, to be separated. Item response theory puts the difficulty of items on the same theta scale as the ability estimate of the test-taker. This means that after each item is taken, the computer adaptive algorithm can compare the current estimate of the participant's ability with the difficulty of the remaining items in the item bank. The CAT algorithm can then choose the item that is closest in difficulty to the test-taker's theta. As long as the item bank is big enough, when a candidate answers correctly then the next item will be harder.

The item bank and parameters of the discounting task items were calibrated independently for each time delay, based on the two parameter logistic (2PL) model. Item responses for the fifteen items within each time delay were then used to simulate a computer adaptive test. The test started at the mid-point (based on difficulty estimates and assuming a Bayesian prior probability distribution) of the distribution of items within each time delay. After each choice, the participant's score was recalculated using maximum-likelihood estimation (ML),[2] and then the next item was selected based on the principle of selecting the most informative item among the remaining items in the item bank, based on the Fisher information function. Following seven forced-stop points (multiples of two from 4–14 items), the test ended. The number of questions was held constant across the two methods being compared in the simulation. To make the binary search method comparable to the IRT method, an additional set of simulations was run on the binary method, using fixed stopping points (multiples of two from 2–8 items).

Simulations

[2] As ML is currently one of the most standard methods used to estimate IRT models and is based on sample data (Bock & Aitkin, 1981; Chen, Hou, & Dodd, 1998; Forero & Maydeu-Olivares, 2009; T. Y. Wang & Vispoel, 1998), we used it as a reference in simulating a computer adaptive test.

Simulations of the 6 time delays (with 5-7 items) found correlations between r = 0.985 and .999 between the newly developed binary search-type algorithm and the standard questionnaire method, indicating the potential for high convergent validity with just half the number of items. The results of the simulation study were correlated with the original dataset and revealed that the binary search-type method was more accurate within a small range of items than the IRT-based method. For six items, the correlations between the responses of the binary search method and the original data ranged from r = 0.83–0.99, depending on the number of questions (i.e. iterations); correlations ranged from r = 0.72–0.94 for responses from the IRT method. The method comparable to IRT, is called Binary or Fixed (fixed referring to the number of questions) – see Figure 1.

Discussion

A simulation study showed that a binary search-like method was more appropriate than IRT-based computer adaptive testing as the small size of the available item bank (15 items per delay) likely influences the accuracy in estimating IRT model parameters. The correlation between original and simulated scores is consistently lower for the IRT method than for non IRT methods (see Figure 1) – probably because of the relatively small item bank available (15 items). Past research shows that sample size affects the stability and accuracy of IRT model parameters, with the magnitude of the variation between sample estimates decreasing as sample size increases (De Ayala, 1999; Wang & Chen, 2005). Following a typical 7:1 or 10:1 ratio (item bank vs. number of items administered), the item bank should consist of ~50-70 items if the maximum number of questions/iterations should not exceed 7.

Both methods require previously gathered sample data, from which to derive the item bank and/or the percentiles. The generalizability of the results obtained using this method therefore depends on the quality of the original sample, although the methodology itself is applicable across multiple contexts. Further, it is important to note that the binary search-type algorithm can yield a finer set of responses than the traditional IRT algorithm, based on the item bank itself as it essentially uses all the items.[3] Thus, experimenters can decide in advance how accurate they want their measurement to be, and accordingly set a stopping rule to meet the criterion.

Tweaking the Initial Computerised Task

Based on the results of the simulation studies, the binary search algorithm was modified by employing a calibration period, followed by a prediction of the discount rate, for each time delay. The calibration period allows for a quicker approach to the general area of the expected percentile in case the subject's predicted percentile is far off, by approaching it in steps of the average standard deviation (in this case, 16) of the sample. Once a preference reversal was indicated, the steps of a simple binary search (half-interval method that was initially used) were performed until the participant reached a single percentile. While for the first timeframe, the starting item was that of the 50th percentile, subsequent timeframes used the percentile of the last item from the previous time delay as a starting point.

Study 2: Validation of the computerised task

The computerised delay discounting task was validated using two independent samples, against a standard measure of delay discounting and other constructs that have been theorised and shown by past research to be related to discounting.

Method

Participants and procedures

[3] An item bank consisting of 10 items per delay allows for only 10 possible switching points, while an item bank that consists of 30 items per time delay, allows for a much wider range of indifference values.

Sample 1: Data were collected online via Amazon's Mechanical Turk. Workers located in the United States received $1 to respond to a set of questionnaires. Participants were omitted for incomplete data or poor quality of responses. 269 participants between 18–80 years of age (*M* = 36.31, SD = 12.68; 117 males) were used in our analyses.

Sample 2: Data were collected online for this study via recruitment advertisements on Amazon's Mechanical Turk application. Workers located in the United States received $0.90 to respond to a set of questionnaires. 313 participants between 18–76 years of age (*M* = 36.14, SD = 13.97; 117 males) were included in our analyses.

Measures of delay discounting

Participants answered the computerised task and the standard measure of delay discounting as the initial or final measure in a battery of measures. Both versions consisted of the same time delays and delayed amounts.

Sample 1: The standard measure was based on past research (Bickel et al., 1999; Rachlin et al., 1991; Stillwell & Tunney, 2012). Four sets of 15 questions were presented in a randomized order to each participant. The immediate reward amounts were $1000, $950, $900, $850, $750, $600, $500, $400, $250, $150, $100, $60, $20, $10 and $1, and time delays were 1 month, 6 months and 5 years. All these amounts and time delays were compared to $1000 at the future time point while all amounts (proportionately less; i.e. $100, $95… $0.1) were also compared to $100 in 1 month – in order to examine the magnitude effect in discounting.[4] The rate of delay discounting was calculated as parameter *k* using a

[4] This repeated measures method controls for different estimates of discounting due to delay length and amount (Chapman, 1996; Chapman & Elstein, 1995; Green, Fristoe, et al., 1994; Green, Fry, et al., 1994; Kirby, 1997; Lane et al., 2003; Mahalingam et al., 2014; Raineri & Rachlin, 1993; Stillwell & Tunney, 2012). This would be infeasible if using real outcomes (Frederick, Loewenstein, & O'Donoghue, 2002).

hyperbolic discounting function.[5] Further, log transformation (to base 10) was used to normalize the data.

Sample 2: The standard measure here was similar to that described above; however, only ten amounts were used as immediate rewards \$1000, \$900, \$750, \$600, \$400, \$250, \$100, \$60, \$20 and \$1; thus, making this version slightly different from the comparison computerised task.

Additional measures used in both samples

The Barratt Impulsiveness Scale (Patton, Stanford, & Barratt, 1995) is a popular 30-item tool that captures the multifaceted nature of impulsivity in the underlying factor structure – three second order factors (attentional, motor and nonplanning impulsiveness) and six oblique first order factors (attention, cognitive instability, motor, perseverance, self-control and cognitive complexity). Internal consistency ranges from Cronbach's $\alpha = 0.83 – 0.27$, while test-retest reliability at one month ranges from Spearman's $\rho = 0.83 – 0.23$ (Stanford et al., 2009). Convergent validity with other self-report measures, including Zuckerman Sensation-Seeking Scale (SSS-V) and Eysenck Impulsiveness Scale (I_7), and behavioral measures, including IMT ratio and DMT, has also been established (Stanford et al., 2009).

All participants responded to the 50-item International Personality Item Pool measure of the NEO Big Five personality traits (Goldberg et al., 2006). Internal consistency ranges between Cronbach's $\alpha = 0.87 – 0.79$, while convergent validity with original factor markers ranged between Pearson's $r = 0.9 – 0.66$ when corrected for unreliability. Reliabilities of the factor markers were assumed to be the same as those of their corresponding IPIP scales.

[5] Preliminary analyses showed that a hyperbolic, time inconsistent function fit the data better than an exponential, time consistent function.

Participants also provided basic demographic information including a question on smoking behavior (‘Do you smoke?’) to which they responded on a 3-point scale (‘Never’, ‘Less than daily’, ‘Daily or more’).

Results

Data analysis approach

Hierarchical linear modelling (HLM) was used to account for multiple observations from the same user (Raudenbush & Bryk, 2002). The delay lengths and delayed amounts were considered interdependent (level 1) compared to smoking behavior, impulsivity, personality factors and demographics that were measured only once (level 2). All continuous variables were centred (Aiken & West, 1991) to minimize multicolinearity.

In our analyses, we controlled for age and gender to rule out important covariates of impulsiveness, personality and smoking behavior, and delay length and reward magnitude to account for the variance within individuals. However, the effects remained highly similar when these covariates were not included. All factors in each model were entered as simultaneous predictors to examine their unique effects.

Internal consistency

Sample 1: Correlations across time delays and delayed amounts of the computerised delay discounting task ranged between r = 0.514–0.716, indicating good internal consistency (see Table 1a). In comparison, correlations across time delays and delayed amounts ranged between r = 0.448–0.731 for the standard measure of delay discounting.

Sample 2: Correlations across time delays and delayed amounts of the computerised delay discounting task ranged between r = 0.488 – 0.687, indicating good internal consistency (see Table 1b). In comparison, correlations across time delays and delayed amounts ranged between r = 0.433–0.703 for the standard measure of delay discounting.

Concurrent validity with a questionnaire measure of delay discounting

Sample 1: The computerised task was correlated with the standard questionnaire measure of delay discounting to establish concurrent validity (see Table 1a). Correlations between each time delay and magnitude ranged from r = 0.492–0.849, with 80% of coefficients being r > 0.6. The correlation coefficient for mean discounting rate (delayed amount of $1000) across the two measures of delay discounting was r = 0.901.

Participants completed 60 items each in the standard measure of delay discounting, while the average number of items per participant was M = 25.98 (SD = 3.44; i.e. 6.5 items per block of items for each length of delay and delayed amount) during the computerised task (see Figure 2a).

Sample 2: The computerised task was correlated with the standard questionnaire measure of delay discounting to establish concurrent validity (see Table 1b). Correlations between each time delay and magnitude ranged from r = 0.429–0.903, with 75% of coefficients being r > 0.6. The correlation coefficient for mean discounting rate (delayed amount of $1000) across the two measures of delay discounting was r = 0.828.

Participants completed 40 items each in the standard measure of delay discounting, while the average number of items per participant was M = 26.6 (SD = 2.88; i.e. 6.65 items per block of items for each length of delay and delayed amount) during the computerised task (see Figure 2b).

Convergent and discriminant validity with the delay effect, magnitude effect and smoking behavior

We tested whether the effects of reward magnitude, delay length and smoking behavior on discounting rates were consistent across the two measures of delay discounting.

Sample 1: Smoking behavior significantly predicted delay discounting regardless of the measure used; people who smoked more often were more impatient for future rewards (see Table 2). These findings are in accordance with past research (Bickel et al., 1999; Daugherty & Brase, 2010; Friedel, DeHart, Madden, & Odum, 2014; Mahalingam et al., 2014).

Table 2b shows the logit model discriminating between regular smokers and social or non-smokers based on delay discounting. For every unit change in delay discounting, the log odds of problematic smoking behavior (versus social smoking or non-smoking) increased by 0.128 (as measured by the computerised task) and 0.117 (as measured by the standard measure). The effects of delay discounting on the odds of being a regular smoker were 1.14 (as measured by the computerised task) and 2.2 (as measured by the standard measure).

Sample 2: Here again, smoking behavior significantly predicted delay discounting regardless of the measure used; people who smoked more often were more impatient for future rewards than non-smokers (see Table 2c), in accordance with Sample 1 and past research.

Convergent validity with personality

We tested whether the effects of Big Five personality traits on discounting rates were consistent across the two measures of delay discounting.

Sample 1: None of the personality traits, except extraversion, significantly predicted discounting; more extraverted people were more impatient with future rewards (see Table 3a). Considering Mahalingam et al. (2014) found relatively small effects between personality traits and discounting in a large international sample[6], it is likely that the smaller sample size here is not conducive to identifying similar effects.

[6] *N* = 5,888; extraversion had the strongest effect size.

Sample 2: Only conscientiousness, as measured by the questionnaire measure, significantly predicted discounting; less conscientious people were more impatient with future rewards (see Table 3b). Here again, it is likely that the relatively small (level 2 N = 309) sample size is not conducive to identifying similar effects.

Convergent validity with impulsiveness

The final goal of Study 2 was to test whether the relationship between different facets of impulsivity, measured by the Barratt Impulsiveness Scale, and delay discounting was consistent across the two measures of discounting in both samples.

Sample 1: Only the cognitive complexity factor was significantly related to delay discounting, irrespective of the measure used (see Table 4a). As cognitive complexity increases, individuals are more impatient for delayed rewards. Other factors of impulsivity did not significantly influence discounting.

Sample 2: Again, only the cognitive complexity factor was significantly related to delay discounting, irrespective of the measure used. As cognitive complexity increases, individuals are more impatient for delayed rewards (see Table 4b). Other factors of impulsivity did not significantly influence discounting, although the motor factor was marginally significant across the two measures.

Discussion

The present study validated a new computerised delay discounting task by showing concurrent validity with standard measures of delay discounting, and convergent validity with addictive behavior, the BIS-11 questionnaire measure of impulsivity and the 50-item IPIP measure of personality; results were relatively consistent across two independent samples that were used. Results were overall supportive of the computerised task with internal consistency/correlations across time delays and delayed amounts within the computerised task ranging between $r = 0.514 - 0.716$ (*Sample 1*) and $r = 0.488 - 0.687$ (*Sample 2*), between the two measures ranging between $r = 0.492 - 0.849$ (*Sample 1*) and $r = 0.429 - 0.806$ (*Sample 2*) and mean discounting rates (for delayed amount = $1000) having a correlation of $r = 0.901$ (*Sample 1*) and $r = 0.828$ (*Sample 2*). Importantly, *Sample 1* participants responded to 26 items ($M = 25.98$, SD = 3.44) on average (for $1000 at three time points and $100 at one time point) during the computerised task – 44% of the number of items they answered in the standard measure consisting of 60 items; while, *Sample 2* participants responded to 27 items ($M = 26.6$, SD = 2.88) on average (for $1000 at three time points and $100 at one time point) during the computerised task – 68% of the number of items they answered in the standard measure consisting of 40 items. Thus, such a significant reduction in the items administered can reduce administration time and related participant inattention or fatigue. These findings are promising in the context of delay discounting, where there are no computerised tasks optimised across reward magnitudes and delays; and in the broader context of computer adaptive testing, where the intention is to develop psychometrically-sound and efficient measures that can be successfully used across clinical and research contexts alike.

In accordance with past research, smoking behavior consistently predicted discounting, regardless of the discounting measure used and across the two samples; daily smokers discount the future more than non-smokers, with discounting rates increasing as smoking behavior increased (Reynolds, Richards, Horn, & Karraker, 2004). Further, we were able to predict smoking status from individuals' discounting behaviour, providing evidence of discriminant validity. Big five personality factors, except extraversion in *Sample 1* and conscientiousness in *Sample 2,* consistently did not predict discounting across the two measures. Though not in accordance with past research on personality and delay discounting (Hirsh et al., 2008; Mahalingam et al., 2014; Ostaszewski, 1996), this is expected as

Mahalingam et al. (2014) found relatively small effect sizes in a large study (n > 5,800) exploring the effects of personality and reward magnitude on discounting. Finally, with impulsivity, only the cognitive complexity factor was related to delay discounting, across the two measures of discounting in both samples. Other factors were unrelated to delay discounting across the two measures, in accordance with Reynolds et al. (2006).

Correlations between the computerised discounting task and the standard questionnaire measure are generally low (e.g. see Figure 3) – in accordance with the view that discounting is a behavioural construct rather than one that can be measured purely by questionnaire methods.

Study 3: Convergent Validity of the Computerised Task

This study further validated the computerised delay discounting task, using three independent samples, against measures of time perspective, survival probability, satisfaction with life, personality, interpersonal trust and discounting of real rewards.

Method

Participants and procedures

Sample 3: Data were collected online for this study via recruitment advertisements on Amazon Mechanical Turk. Workers received $2 to respond to a set of questionnaires. 189 participants between 18–72 years of age (*M* = 36.16, SD = 12.11; 108 males) and located in the United States were included in these analyses.

Sample 4: Data were collected online for this study via recruitment advertisements on Amazon's Mechanical Turk. Workers located in the United States received $0.90 to respond to a set of questionnaires. 208 participants between 18-74 years of age (M = 38.01, SD = 13.42; 75 males) were included in these analyses.

Sample 5: Data were collected online for this study via recruitment advertisements on university and departmental mailing lists and Facebook posts. Participants responded to a set of questionnaires on a voluntary basis and received no remuneration for the same. 151 participants between 18-79 years of age (M = 27.81, SD = 18.71; 72 males) were included in these analyses.

Sample 6: Data were collected online via Amazon's Mechanical Turk. Workers located in the United States received $0.90 to respond to a set of questionnaires. 420 participants between 18-74 years of age (M = 35.79, SD = 11.5; 144 males) were included in these analyses.

Measures used

All samples engaged in the computerised delay discounting task and provided demographic information. Participants also responded to additional measures for the purpose of unrelated research.

Sample 3: Participants also responded to the Zimbardo Time Perspective Inventory (ZTPI; Zimbardo & Boyd, 1999) and the possibility of winning a lottery of either a $60 prize paid immediately or a $100 prize paid in three months' time, to measure discounting of real rewards. The item was scored such that higher scores imply preference for immediate rewards and lower scores imply a preference for delayed rewards, as in the computerised task's log(k) value.

The ZTPI (Zimbardo & Boyd, 1999) is a self-report measure consisting of 56 items addressing attitudes and behaviors relating to time perspective and a 5-point Likert scale. It consists of five scales distinguished on the basis of factor analysis: Past-Negative (α = .83), Present Hedonistic (α = .81), Future (α = .75), Past Positive (α = .62) and Present Fatalistic (α = .73). The scales have also revealed adequate internal consistencies (in the range of .63 to .84) across numerous cultural contexts (Sircova & Mitina, 2008).

Sample 4: Participants also responded to the 50-item International Personality Item Pool (Goldberg et al., 2006), Subjective Probability of Survival (Chao, Szrek, Pereira, & Pauly, 2009; Smith, Taylor, & Sloan, 2001) and Satisfaction With Life Scale (Diener, Emmons, Larsen, & Griffin, 1985).

SP (Chao et al., 2009; Smith et al., 2001) was measured using two items asking participants to estimate the likelihood of them surviving a specific period of time – 1 year and 25 years. The two items were aggregated to form a composite score.

The SWL is a brief, well-established, 5-item scale that measures global mental judgments of life satisfaction in the general population (Diener et al., 1985). It is a psychometrically sound measure with high internal consistency (.87) and test-retest reliability (.82 at two months); criterion-related, discriminant and convergent validity have also been established by the authors.

Sample 5: Participants also responded to the SP (Chao et al., 2009; Smith et al., 2001), Rotter's Interpersonal Trust (Rotter, 1967) and the 20-item International Personality Item Pool (Goldberg et al., 2006).

Sample 6: Participants also responded to the 50-item International Personality Item Pool (Goldberg et al., 2006).

Results

Multiple correlations were used to analyse the data, for ease of comparing multiple measures across samples.

Internal consistency

The computerised delay discounting task showed moderate-high internal consistency across the five independent samples (see Table 6-9).

Sample 3: Correlations across time delays and delayed amounts of the computerised delay discounting task ranged between r = 0.138 - 0.454, indicating moderate internal consistency (see Table 6).[7]

Sample 4: Correlations across time delays and delayed amounts of the computerised delay discounting task ranged between r = 0.412 - 0.820, indicating moderate-high internal consistency (see Table 7). The average number of items per participant was M = 22.01 (SD = 7.39; i.e. 6 items per block of items for each length of delay and delayed amount) during the computerised task (see Figure 4a) – in this case, 63% shorter than the equivalent standard questionnaire measure.

Sample 5: Correlations across time delays and delayed amounts of the computerised delay discounting task ranged between r = 0.370 - 0.727, indicating moderate-high internal consistency (see Table 8). The average number of items per participant was M = 27.31 (SD = 4.09; i.e. 7 items per block of items for each length of delay and delayed amount) during the computerised task (see Figure 4b) – in this case, 55% shorter than the equivalent standard questionnaire measure.

Sample 6: Correlations across time delays and delayed amounts of the computerised delay discounting task ranged between r = 0.502 - 0.717, indicating moderate-high internal consistency (see Table 9). The average number of items per participant was M = 27.4 (SD = 3.98; i.e. 7 items per block of items for each length of delay and delayed amount) during the computerised task (see Figure 4c) – in this case, 55% shorter than the equivalent standard questionnaire measure.

[7] Item-level data was not available for this sample to provide the average test length per participant.

Concurrent validity with discounting of real rewards

Sample 3: The computerised task was correlated with an item measuring discounting of real rewards. The computerised task using hypothetical rewards showed a correlation of r = 0.601 with the item measuring discounting of real rewards (see Table 6).

Convergent validity with personality

Considering Mahalingam et al. (2014) found relatively small effects between personality traits and discounting in a large international sample[8], it is likely that the relatively small sample sizes here are not conducive to identifying similar effects, as observed in the previous validation study. However, the samples in the current study are relatively consistent (see Table 7-9) with the findings from Study 2.

Sample 4: Only conscientiousness was significantly correlated with mean discounting behavior, as in Sample 2. The more conscientious an individual, the more impatient they tended to be for immediate outcomes (see Table 7).

Sample 5: As in Sample 1, only extraversion was marginally related to mean discounting behavior – significance varying across delays and reward magnitudes. Extroverted individuals tended to be more impatient for delayed outcomes (see Table 8).

Sample 6: Here, consistently, none of the five personality factors were correlated with mean discounting rates (see Table 9).

Convergent validity with survival probability

Sample 4: Survival probability was not significantly correlated with mean discounting behavior, although there was a significant relationship with discounting at 6 months for a $1000 delayed reward and at 1 month for a $100 delayed reward (see Table 7).

[8] N = 5,888; extraversion had the strongest effect size.

Sample 5: Here again, survival probability was not significantly correlated with mean discounting behavior, although there was a marginal significant relationship with mean discounting, at 1 month for a $1000 delayed reward and at 1 year for a $1000 delayed reward (see Table 8).

Convergent validity with interpersonal trust

Sample 5: Interpersonal trust was marginally correlated with discounting behavior, across delays and reward magnitudes, in accordance with Michaelson et al. (2013). As the level of interpersonal trust increased, individuals were less impatient for delayed outcomes and able to wait (see Table 8).

Divergent validity with time perspective

Sample 3: Only the present-fatalistic factor was significantly related to delay discounting. As present-fatalistic scores increase, individuals are more impatient for delayed rewards (see Table 6). Other factors of time perspective were not significantly correlated with discounting, primarily in accordance with Stolarski, Bitner & Zimbardo (2011) who concluded that isolated time perspective dimensions may not explain the tendency to delay gratification regardless of theoretical relevance.

Divergent validity with life satisfaction

Sample 4: SWL was not significantly correlated with discounting behavior, across delays and reward magnitudes (see Table 7).

Divergent validity with age and gender

Delay discounting as measured by the computerised task was found to be consistently unrelated to age and gender across all samples (see Table 6-9), in accordance with Study 2 and in partial accordance with inconclusive existing research (Mahalingam et al., 2014; Reynolds, Karraker, Horn, & Richards, 2003; Reynolds et al., 2004).

Discussion

The present study further validated the newly developed computerised delay discounting task by showing concurrent validity with discounting of real rewards; convergent validity with the 20-item and 50-item IPIP measures of personality, Rotter's Interpersonal Trust Scale and Subjective Probability of Survival; and divergent validity with Satisfaction with Life Scale, Zimbardo's Time Perspective Inventory and age and gender.

Results were overall supportive of the computerised task with correlations across time delays and delayed amounts within the computerised task ranging between r = 0.138–0.820. Importantly, participants in Sample 4-6 responded to approximately 25 items (M = 22.01-27.4, SD = 3.98-7.39) on average (for $1000 at three time points and $100 at one time point) during the computerised task – 37-45% of the number of items they would have answered in an equivalent standard measure consisting of 60 items. Here again, this amounts to less than 7 items per delay length. As the total number of items to be administered increases, especially the number of immediate amounts, the proportionate difference between the standard measure and the computerised task will also increase. Thus, such a significant reduction in the items administered can reduce administration time and related participant inattention or fatigue. These findings are in accordance with Study 2. Finally, the computerised task also showed a correlation of r = 0.601 with discounting of real rewards, providing additional evidence of concurrent validity.

The computerised task also showed evidence of convergent validity. Big five personality factors, except extraversion and conscientiousness, consistently did not predict discounting behavior across the samples. Though not entirely in accordance with past research on personality and discounting behavior (Hirsh et al., 2008; Mahalingam et al., 2014; Ostaszewski, 1996), this is expected as Mahalingam et al. (2014) found relatively small effect sizes in a large study (n > 5,800) exploring the effects of personality and reward magnitude on discounting behavior. Across Sample 4 and 5, survival probability was not significantly correlated with mean discounting, although the construct was correlated at times with individual time delays.

The computerised task also showed evidence of divergent validity across the samples. In Sample 3, only the present-fatalistic dimension of ZTPI was correlated with discounting behavior, mostly in accordance with Stolarski, Bitner & Zimbardo (2011) who concluded that isolated time perspective dimensions may not explain the tendency to delay gratification regardless of theoretical relevance. Similarly, discounting behavior was not correlated with SWL in Sample 4. Finally, age and gender were consistently unrelated to delay discounting in partial accordance with previous research (Mahalingam et al., 2014; Reynolds et al., 2003; Reynolds et al., 2004) and consistent with Study 2.

Correlations between the computerised discounting task and the standard questionnaire measure are generally low (e.g. see Figure 3) – in accordance with the view that discounting is a behavioural construct rather than one that can be measured purely by questionnaire methods.

Limitations and Future Directions

Despite the value of an adaptive measure of delay discounting, certain limitations should not be overlooked. First, as in traditional psychometric tests, the validity of the measure is dependent on the normative data used. Researchers should consider the population under study and its similarity to the myPersonality dataset (Stillwell & Kosinski, 2011) when using the norms we provide (Table 5). However, the methodological approach is relevant across situations and can be adopted universally. Second, for additional time delays or delayed amounts, traditional questionnaire measures will still need to be used until sufficient data has been collected to calculate percentiles.

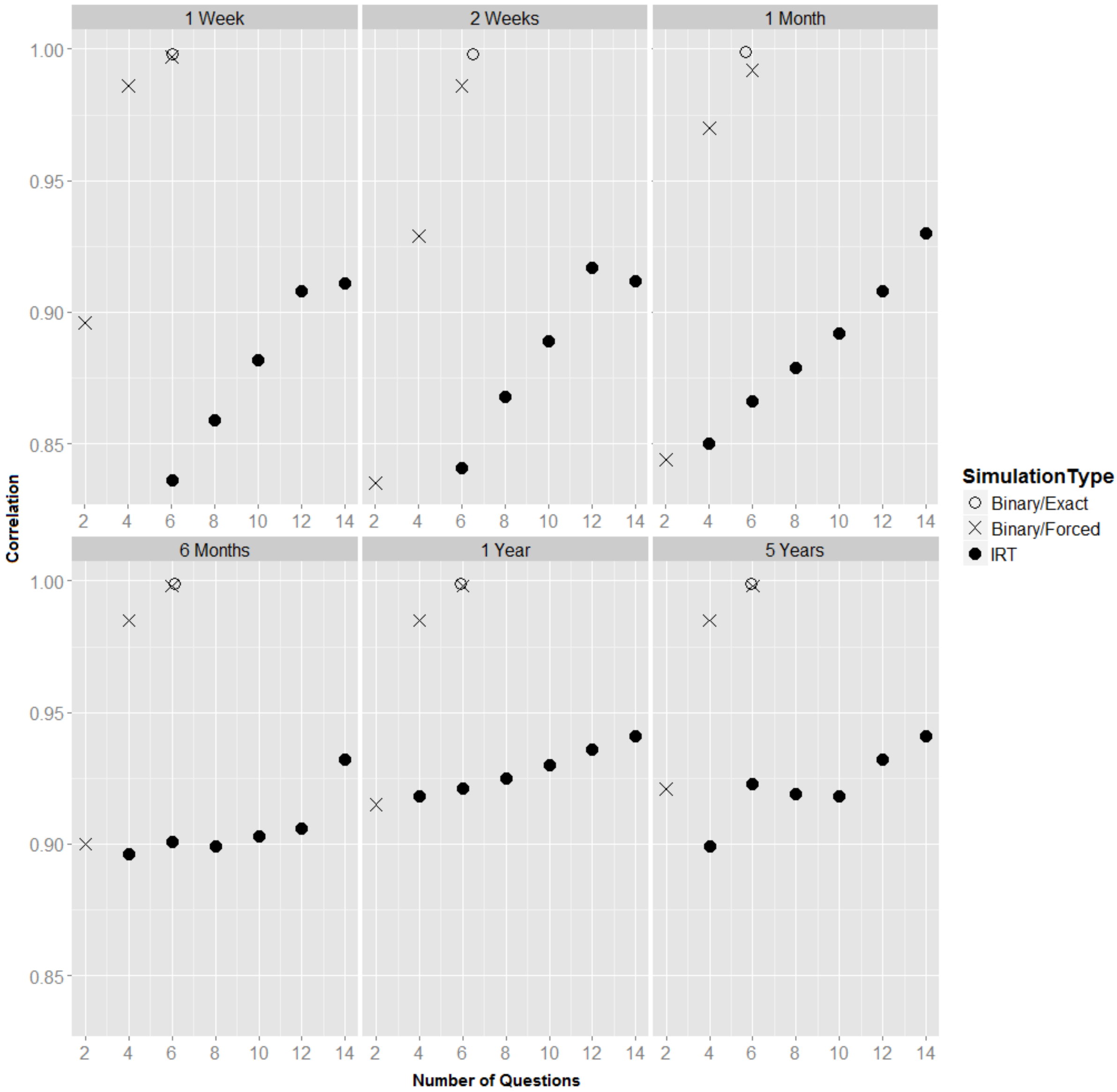
Further research is also required to explore implementation of this CAT to alternative mathematical models or to the atheoretical AUC method of modelling delay discounting.

Conclusion

The objective of this research was to develop and validate a psychometrically sound and efficient computer adaptive measure of delay discounting using a large dataset of *N* = 4,190 participants. First, a binary search-type algorithm was developed to measure delay discounting. Next, across six samples (N = 1550) the computerised task showed evidence of concurrent validity with two standard measures of delay discounting and an item measuring discounting of real rewards; convergent validity with smoking behaviour, the BIS-11 questionnaire measure of impulsivity, Rotter's Interpersonal Trust inventory, Subjective probability of Survival scale, and 20-item and 50-item IPIP measures of personality; and divergent validity with Zimbardo's Time Perspective Inventory, Satisfaction With Life Scale, and age and gender. The computerised task was more effective than the standard measure in identifying the relationship between smoking behavior and delay discounting, showing evidence of discriminant validity. The task was 55-63% shorter than the 60-item full-length measure, across five independent samples; thereby significantly reducing administration time and participant fatigue or inattention. Finally, the task includes a range of time delays and can also be applied to other reward magnitudes. In conclusion, the computerised task is a psychometrically sound and efficient measure of delay discounting that can be universally adopted by researchers and clinicians alike.

List of Figures

Figure 1 Relationship between number of questions and correlation of simulation and real data for different timeframes in the $1000 condition



[a] *For the sake of readability, two data points (Correlation < .80, 1 & 2 week condition) were omitted to increase readability of the graphs.*

Figure 2a-b Average number of items answered per time point per participant in Sample 1-2 during the computerised task

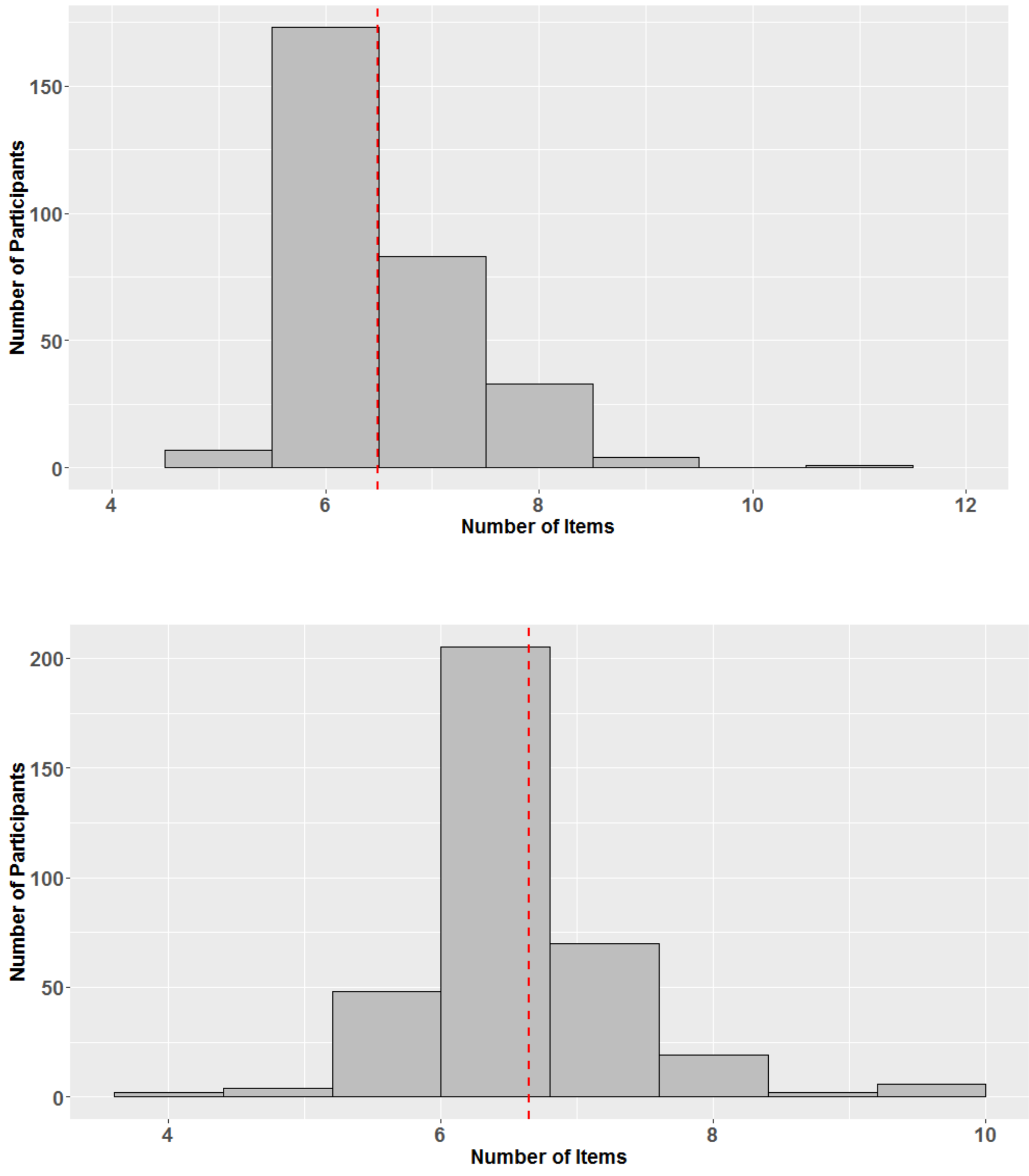


Figure 3 shows the correlation between prediction models from the computerised task and the standard measure in Sample 1

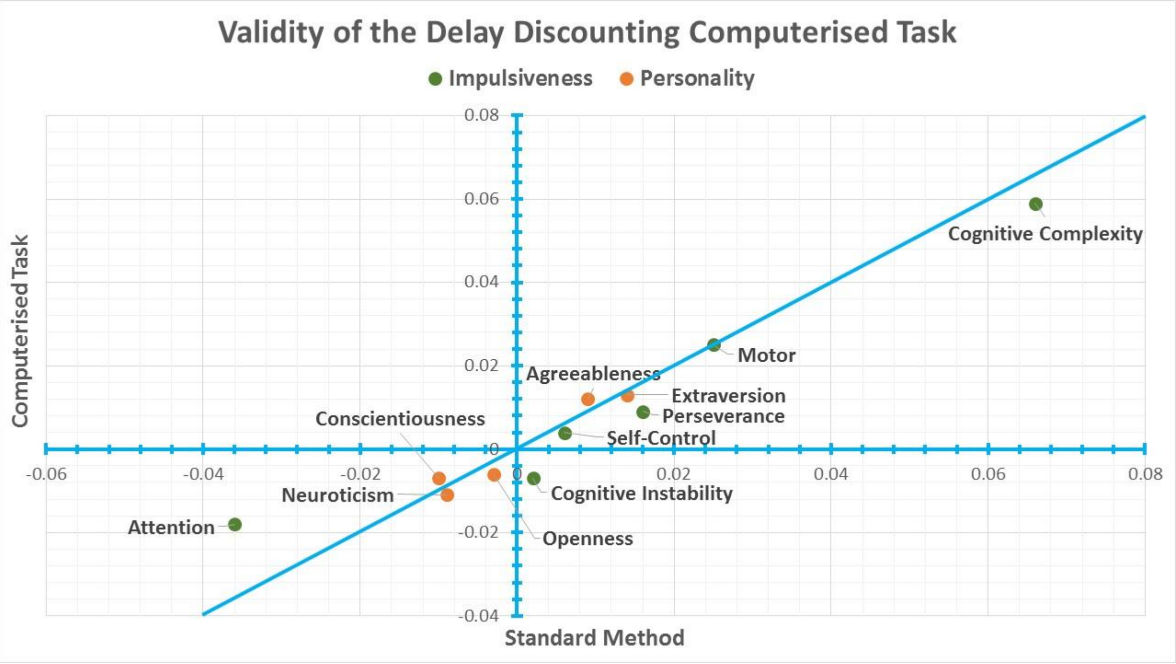


Table 1a Correlations between the computerised task and standard measure of delay discounting in Sample 1

	Predictors	Computerised task					Standard measure			
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Computerised task	Mean discounting (amount: $1000)	-								
	Time: 1 month; amount: $1000	.841	-							
	Time: 6 months; amount: $1000	.915	.689	-						
	Time: 5 years; amount: $1000	.854	.514	.716	-					
	Time: 1 month; Amount: $100	.742	.711	.642	.527	-				
Standard measure	Mean discounting (amount: $1000)	.901	.741	.808	.796	.701	-			
	Time: 1 month; amount: $1000	.755	.773	.644	.500	.685	.812	-		
	Time: 6 months; amount: $1000	.809	.677	.752	.670	.632	.920	.698	-	
	Time: 5 years; amount: $1000	.765	.492	.672	.849	.510	.84	.448	.661	-
	Time: 1 month; Amount: $100	.698	.660	.592	.566	.640	.713	.731	.627	.508

[a] *Time delays - 1 month, 6 months, and 5 years; delayed amounts - $1000, $100.*

Table 1b Correlations between the computerised task and standard measure for measuring delay discounting in Sample 2

	Predictors	Computerised task					Standard measure			
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Computerised task	Mean discounting (amount: $1000)	-								
	Time: 1 month; amount: $1000	.835	-							
	Time: 6 months; amount: $1000	.891	.663	-						
	Time: 5 years; amount: $1000	.85	.498	.687	-					
	Time: 1 month; Amount: $100	.657	.63	.604	.488	-				
Standard measure	Mean discounting (amount: $1000)	.828	.662	.795	.702	.61	-			
	Time: 1 month; amount: $1000	.673	.657	.63	.458	.612	.793	-		
	Time: 6 months; amount: $1000	.773	.623	.806	.636	.509	.903	.663	-	
	Time: 5 years; amount: $1000	.679	.465	.634	.716	.447	.848	.433	.681	-
	Time: 1 month; Amount: $100	.635	.656	.636	.429	.683	.69	.703	.62	.467

**Time delays: 1 month, 6 months, 5 years; delayed amounts: $1000, $100.*

Table 2a Effects of reward magnitude, delay length and smoking behaviour on delay discounting in Sample 1

Predictors	Computerised task			Standard measure		
	b	**t**	**p-value**	***b***	**t**	**p-value**
Level 1						
Time	-.007	-10.41	< .001	.-.011	-17.80	< .001
Amount	-.308	-7.78	< .001	-1.432	-41.46	< .001
Level 2						
Smoking behaviour	.228	4.06	< .001	.15	2.97	< .001
Age	-.002	-.65	.519	-.004	-1.21	.228
Gender	.117	1.33	.186	.128	1.61	.108

[a] *All numbers are unstandardized regression coefficients.* [b] *Computerised task: Level 1 N = 1063, level 2 N = 269; standard measure: Level 1 N = 1058, level 2 N = 269.*

Table 2b Discriminating between regular smokers and social or non-smokers from delay discounting in Sample 1

Predictors	Computerised task			Standard measure		
	b	**z**	**p-value**	***b***	**z**	**p-value**
Delay discounting	.932	3.75	< .001	.79	2.96	.003
Age	.014	1.05	.296	.017	1.33	.185
Gender	-.335	-.99	.329	-.268	-.81	.418

[a] *All numbers are unstandardized logit regression coefficients.* [b] *Computerised task N = 258, standard measure N = 253.*

Table 2c Effects of reward magnitude, delay length and smoking status on discounting rates in Sample 2

Predictors	Computerised task				Standard measure			
	b	**t**	**Pseudo R^2**	**p-value**	***b***	**t**	**Pseudo R^2**	**p-value**
Level 1								
Time	-.009	-30.52		< .001	.-.011	-47.08		< .001
Amount	.137	17.11		< .001	.164	23.99		< .001
Level 2								
Smoking status	.162	3.54		< .001	.151	4.03		< .001
Age	.003	.98		.329	.001	.29		.774
Gender	-.041	-.54		.593	.024	.39		.698

**All numbers are unstandardized regression coefficients. Computerised task: Level 1 N = 4880, level 2 N = 309; standard measure: Level 1 N = 4864, level 2 N = 309.*

Table 3a Effects of personality on delay discounting in Sample 1

Predictors	Computerised task			Standard measure		
	b	**t**	**p-value**	***b***	**t**	**p-value**
Level 1						
Time	-.007	-10.41	< .001	-.011	-17.81	< .001
Amount	-.309	7.81	< .001	-1.433	-41.50	< .001
Level 2						
Openness	-.003	-.40	.692	-.006	-0.94	.349
Conscientiousness	-.01	-1.36	.174	-.007	-1.19	.234
Extraversion	.014	2.62	.009	.013	2.70	.007
Agreeableness	.009	1.39	.165	.012	2.08	.038
Neuroticism	-.009	-1.61	.108	-.011	-2.46	.015
Age	.002	.47	.637	-.000	-.00	.999
Gender	.002	-.02	.985	-.003	-.04	.969

[a] *All numbers are unstandardized regression coefficients.* [b] *Computerised task: Level 1 N =1063, level 2 N =269; standard measure: Level 1 N =1058, level 2 N =269.*

Table 3b Effects of personality on discounting rates in Sample 2

Predictors	Computerised task				Standard measure			
	b	**t**	**Pseudo R^2**	**p-value**	***b***	**t**	**Pseudo R^2**	**p-value**
Level 1								
Time	-.009	-30.51		< .001	-.011	-47.07		< .001
Amount	.137	17.11		< .001	.164	23.99		< .001
Level 2								
Openness	.005	.42		.676	.003	0.28		.778
Conscientiousness	-.02	-1.61		.109	-.021	-2.08		.038
Extraversion	-.017	-1.34		.182	-.014	-1.4		.162
Agreeableness	.002	.13		.897	-.011	-1.05		.297
Neuroticism	-.002	-.18		.854	-.003	-.39		.699
Age	.003	.99		.323	.001	.26		.795
Gender	-.076	-.97		.332	-.007	-.11		.914

**All numbers are unstandardized regression coefficients. Computerised task: Level 1 N = 4880, level 2 N = 309; standard measure: Level 1 N = 4864, level 2 N = 309.*

Table 4a Relationship between impulsiveness and delay discounting in Sample 1

Predictors	Computerised task			Standard measure		
	b	**t**	**p-value**	***b***	**t**	**p-value**
Level 1						
Time	-.007	-10.41	< .001	-.011	-17.81	< .001
Amount	-.309	-7.81	< .001	-1.433	-41.48	< .001
Level 2						
Attention	-.036	-1.81	.072	-.018	-1.01	.315
Motor	.025	1.61	.110	.025	1.80	.073
Cognitive Instability	.002	.06	.953	-.007	-.25	.803
Perseverance	.016	.60	.548	.009	.36	.700
Self-Control	.006	.36	.723	.004	.26	.793
Cognitive Complexity	.066	3.32	< .001	.059	3.37	< .001
Age	-.002	-.41	.681	-.0003	-.09	.358
Gender	.01	1.15	.252	.118	1.51	.132

[a] *All numbers are unstandardized regression coefficients.* [b] *Computerised task: Level 1 N = 1063, level 2 N = 269; standard measure: Level 1 N = 1058, level 2 N = 269.*

Table 4b Relationship between impulsiveness and discounting behaviour in Sample 2

Predictors	Computerised task				Standard measure			
	b	**t**	**Pseudo R^2**	**p-value**	***b***	**t**	**Pseudo R^2**	**p-value**
Level 1								
Time	-.009	-30.51		< .001	-.011	-47.07		< .001
Amount	.137	17.11		< .001	.164	24		< .001
Level 2								
Attention	-.022	-1.25		.213	-.026	-1.8		.072
Motor	.026	1.84		.067	.023	1.97		.05
Cognitive Instability	.011	.41		.682	-.006	-.3		.765
Perseverance	-.02	-.89		.372	-.012	-.64		.522
Self-Control	-.01	-.63		.529	-.016	-.13		.897
Cognitive Complexity	.062	3.75		< .001	.05	3.64		< .001
Age	.003	.94		.35	.0004	.19		.849
Gender	-.081	-1.07		.286	-.012	-.18		.854

**All numbers are unstandardized regression coefficients. Computerised task: Level 1 N =4880, level 2 N =309; standard measure: Level 1 N =4864, level 2 N =309.*

Table 5 Percentiles for $1000 and $100 delayed amounts

Delayed amount: $1000										
1 Week	**x0**	**x1**	**x2**	**x3**	**x4**	**x5**	**x6**	**x7**	**x8**	**x9**
00x	1000	994	993.5	993	992.5	992	991.5	991	990.5	990
01x	989.5	989	988.5	988	987.5	987	986.5	986	985.5	985
02x	984.5	984	983.5	983	982.5	982	981.5	981	980.5	980
03x	979.5	979	978.5	978	977	976	975	974	973	972
04x	971	969	968	966	964	963	960	958	955	952
05x	950	947	945	942	940	937	934	930	926	921
06x	916	910	904	899	893	887	881	875	868	860
07x	852	843	832	820	807	793	778	763	747	731
08x	714	696	677	658	638	617	596	573	550	526
09x	501	475	448	421	393	364	335	304	273	242
10x	209									
2 Weeks										
00x	1000	990.5	990	989.5	989	988.5	988	987.5	987	986
01x	985	984.5	984	983	982	981	979	978	977	975
02x	974	973	971	970	968	966	965	963	961	959
03x	957	955	952	950	947	945	942	939	936	933
04x	930	927	923	920	916	912	908	904	900	896
05x	892	888	885	881	876	872	867	862	856	850
06x	843	837	830	823	816	809	801	794	785	776
07x	767	756	745	732	719	705	691	677	661	646
08x	630	613	596	578	560	541	522	502	481	460
09x	438	415	392	368	344	319	294	268	241	214
10x	187									

1 Month										
00x	1000	984	983	982	981	980	979	978	977	976
01x	974	972	971	969	967	965	963	961	958	956
02x	953	951	948	945	942	939	935	932	928	925
03x	921	917	913	909	905	901	897	892	887	883
04x	878	873	869	864	859	854	849	844	838	832
05x	826	820	813	807	800	793	786	779	772	764
06x	756	748	739	730	721	711	702	692	683	673
07x	662	651	640	628	615	601	586	571	556	541
08x	525	508	491	474	456	438	419	399	379	359
09x	338	316	294	272	249	226	202	177	152	127
10x	101									
6 Months										
00x	1000	990	985	980	975	970	965	959	954	949
01x	943	938	932	926	921	915	909	903	897	891
02x	885	878	872	866	859	853	846	840	833	826
03x	819	813	806	799	792	785	777	770	763	756
04x	748	741	733	726	718	710	702	694	686	678
05x	670	661	653	645	636	628	619	611	602	593
06x	584	575	565	555	545	535	525	515	505	494
07x	484	473	461	449	437	425	412	399	385	372
08x	358	344	329	315	300	285	269	253	237	221
09x	205	188	171	153	136	118	100	81	62	43
10x	24									

1 Year	x0	x1	x2	x3	x4	x5	x6	x7	x8	x9
00x	1000	989	981	973	966	958	950	942	934	926
01x	918	910	903	895	887	879	871	863	855	847
02x	839	830	822	814	806	798	790	781	773	765
03x	757	749	740	732	724	716	707	699	690	682
04x	673	665	656	648	639	630	622	613	604	596
05x	587	578	569	561	552	544	535	527	518	509
06x	500	491	481	471	461	451	441	430	420	409
07x	399	388	376	365	353	342	330	317	305	292
08x	280	267	254	241	228	214	200	187	172	158
09x	144	129	114	99	84	69	53	38	22	6
10x	1									

5 Years	x0	x1	x2	x3	x4	x5	x6	x7	x8	x9
00x	1000	959	943	928	913	898	883	869	855	840
01x	826	813	799	786	772	759	746	734	721	709
02x	696	684	672	661	649	637	626	615	605	594
03x	584	574	564	554	545	535	525	516	506	497
04x	488	479	470	461	452	443	435	426	417	409
05x	400	391	382	373	365	356	347	339	330	322
06x	313	305	296	287	278	269	260	251	242	232
07x	223	214	205	196	188	179	171	163	155	146
08x	138	130	122	114	106	99	91	83	76	68
09x	438	415	392	368	344	319	294	268	241	214
10x	1									

Delayed amount: $100										
1 Month										
00x	100	99.5	99	98.5	98	97.5	97	96.5	96	95.5
01x	95	94.5	94	93.5	93	92.5	92	91.5	91	90.5
02x	90	89.5	89	88.5	88	87.5	87	86.5	86	85.5
03x	85	84.5	84	83.5	83	82.5	82	81.5	81	80.5
04x	80	79.5	79	78.5	78	77	76.5	76	75	74.5
05x	74	73	72.5	72	71	70.5	70	69	68.5	68
06x	67	66.5	66	65	64	63.5	63	62	61	60.5
07x	60	59	58	57	56	55	53	52	51	50
08x	48	47	45	44	42	41	39	38	36	34
09x	32	30	29	27	25	23	20	18	16	14
10x	12									

Table 6 Pearson's correlation coefficients (below diagonal) and significance levels (above diagonal) for variables examined in Sample 3

Factors	1	2	3	4	5	6	7	8	9	10	11	12	13
1. DD: 1 mth-$1000	-	2.00E-09	.005	7.00E-07	2.00E-07	0	.107	.132	.92	.093	.13	.77	.68
2. DD: 6 mths-$1000	.417	-	9.00E-08	2.00E-13	.006	1.00E-09	.018	.399	.98	.253	.017	.83	.05
3. DD: 1 yr-$1000	.204	.377	-	2.00E-07	.002	8.00E-05	.061	.468	.267	.969	.254	.84	.08
4. DD: Mean-$1000	-.35	-.5	-.37	-	.325	0	.27	.213	.698	.366	.012	.09	.45
5. DD: 1 mth-$100	.371	.2	.22	-.072	-	.111	.338	.649	.679	.02	.321	.78	.62
6. DD: Real reward	-.27	-.425	-.284	.612	-.116	-	.546	.019	.805	.791	.147	.49	.63
7. ZTPI: Past negative	-.12	-.172	-.136	.081	-.07	.044	-	.107	.049	0	0	.18	.07
8. ZTPI: Present hedonistic	-.11	-.062	-.053	.091	-.033	.17	.118	-	0	.128	0	.07	.95
9. ZTPI: Future	-.01	-.002	-.081	-.028	-.03	-.018	-.14	-.38	-	0	0	.21	.11
1. ZTPI: Past positive	-.12	.084	.003	-.066	-.169	-.019	-.43	.111	.364	-	0	.07	.01
11. ZTPI: Present fatalistic	-.11	-.173	-.083	.183	-.073	.106	.438	.453	-.51	-.34	-	.36	.63
12. Age	.021	-.016	.015	-.124	.021	-.051	-.1	-.13	.093	.133	-.07	-	0
13. Gender (F)	-.03	-.143	-.126	.055	-.036	.035	-.13	-.01	.118	.181	-.04	.25	-

Table 7 Pearson's correlation coefficients (below diagonal) and significance levels (above diagonal) for variables examined in Sample 4

Factors	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. DD: 1 mth-$1000	-	0	3E-12	0	6E-06	.418	.712	.816	.970	.016	.660	.693	.856	.372
2. DD: 6 mths-$1000	.820	-	0	0	1E-10	.036	.268	.095	.570	.043	.527	.307	.497	.795
3. DD: 1 yr-$1000	.631	.785	-	0	3E-05	.584	.948	.412	.910	.046	.278	.522	.381	.660
4. DD: Mean-$1000	.894	.951	.918	-	2E-10	.171	.722	.378	.875	.030	.718	.583	.363	.654
5. DD: 1 mth-$100	.440	.592	.412	.587	-	.045	.463	.026	.615	.338	.498	.004	.016	.967
6. SP	-.083	-.212	-.056	-.139	-.203	-	.900	.298	.161	.239	.545	.490	.001	.001
7. SWL	-.038	-.113	-.007	-.036	-.075	-.013	-	.007	.292	2E-05	.000	.956	.300	.008
8. Extraversion	-.024	-.170	-.084	-.090	-.224	-.106	.271	-	.002	.271	.009	.706	.003	.090
9. Agreeableness	.004	-.058	-.012	-.016	-.051	-.143	.108	.307	-	.078	.821	.000	.308	.020
10. Conscientiousness	.242	.205	.202	.219	.098	-.120	.413	.112	.179	-	.124	.378	.039	.050
11. Neuroticism	-.045	.065	.111	.037	.069	-.062	-.365	-.262	-.023	-.156	-	.228	.104	.326
12. Openness	-.040	-.104	-.066	-.056	-.285	-.070	-.006	-.039	.383	-.090	.123	-	.011	.635
13. Age	.019	.069	.089	.093	.242	-.334	.106	.296	.104	.209	-.165	-.256	-	6E-05
14. Gender	.091	.027	.045	.046	.004	-.333	.266	.172	.235	.198	-.100	-.049	.394	-

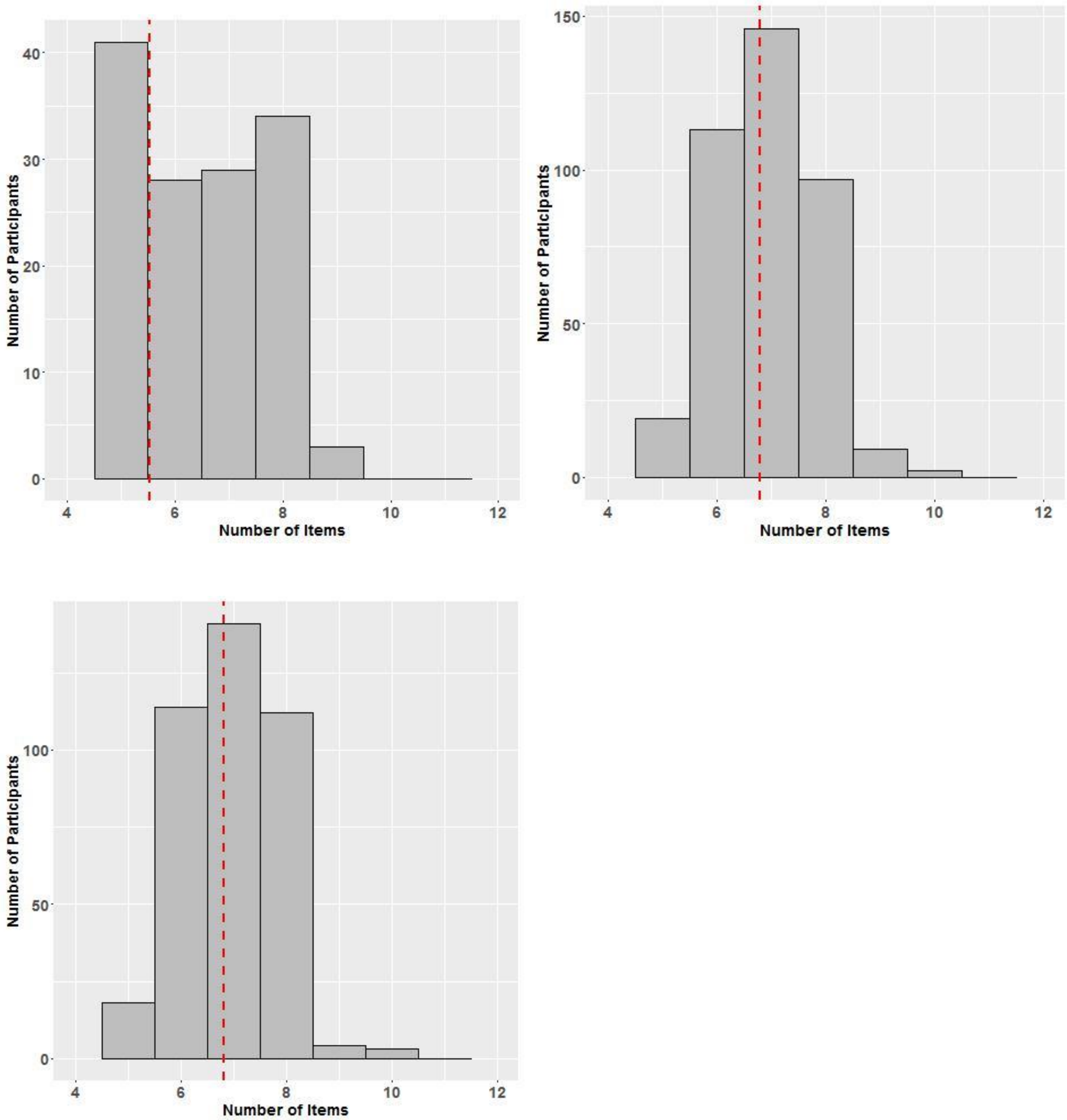
Table 8 Pearson's correlation coefficients (below diagonal) and significance levels (above diagonal) for variables examined in Sample 5

Factors	1	2	3	4	5	6	7	8	9	10	11	12	13
1. DD: 1 mth-$1000	-	0	1E-14	0	0	.536	.061	.646	.834	.084	.169	.964	.933
2. DD: 6 mths-$1000	.727	-	0	0	2.22E-14	.784	.949	.075	.422	.087	.044	.496	.564
3. DD: 1 yr-$1000	.500	.542	-	0	3.89E-08	.052	.061	.037	.371	.301	.042	.710	.140
4. DD: Mean-$1000	.846	.891	.831	-	2.22E-16	.293	.093	.051	.626	.143	.062	.743	.804
5. DD: 1 mth-$100	.573	.497	.370	.529	-	.765	.074	.403	.048	.054	.740	.937	.010
6. Openness	-.043	.019	.135	.073	-.021	-	.788	.002	.001	.229	.191	.092	.679
7. Conscientiousness	.130	-.004	.130	.117	-.124	.019	-	.788	.003	.003	.106	.162	.776
8. Extraversion	.032	.124	.144	.136	.058	.210	-.019	-	.782	.028	.326	.642	.040
9.Agreeableness	-.015	.056	.062	.034	-.137	.235	.206	.019	-	.500	.535	.044	.008
1. Neuroticism	.120	.119	.072	.102	.134	-.084	-.206	-.152	-.047	-	.002	.003	.027
11. Interpersonal trust	-.096	-.140	-.141	-.130	.023	-.091	.112	-.068	-.043	-.209	-	.557	.169
12. Age	.003	.047	-.026	.023	-.006	.117	.097	.032	.140	-.202	.041	-	.349
13. Gender	-.006	-.040	.103	.017	-.178	-.029	.020	-.142	.183	.153	-.096	-.065	-

Table 9 Pearson's correlation coefficients (below diagonal) and significance levels (above diagonal) for variables examined in Sample 6

Factors	1	2	3	4	5	6	7	8	9	10	11	12
1. DD: 1 mth-$1000	-	0	0	0	0	.100	.157	.041	.105	.213	.880	.194
2. DD: 6 mths-$1000	.717	-	0	0	0	.417	.578	.511	.522	.859	.464	.648
3. DD: 1 yr-$1000	.513	.668	-	0	0	.016	.012	.051	.061	.007	.091	.771
4. DD: Mean-$1000	.838	.912	.854	-	0	.980	.774	.759	.918	.463	.413	.466
5. DD: 1 mth-$100	.661	.634	.502	.710	-	.167	.110	.829	.244	.461	.817	.132
6. Openness	.080	.040	-.117	.001	-.068	-	.000	.000	.000	.000	.246	.307
7. Conscientiousness	.069	.027	-.123	-.014	-.078	.805	-	.000	.000	.000	.498	.288
8. Extraversion	.100	.032	-.095	.015	-.011	.690	.631	-	.000	.000	.600	.151
9. Agreeableness	.079	.031	-.091	.005	-.057	.854	.802	.738	-	.000	.513	.011
1. Neuroticism	.061	-.009	-.132	-.036	-.036	.710	.784	.682	.730	-	.052	.432
11. Age	-.007	.036	.083	.040	-.011	-.057	-.033	-.026	.032	.095	-	.270
12. Gender	.064	.022	.014	.036	.074	.050	.052	.070	.124	-.038	-.054	-

Figure 4a-c Average number of items answered per time point per participant in Sample 4-6 (left-right) during the computerised task



References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. London: Sage.

Bickel, W. K., Odum, A. L., & Madden, G. J. (1999). Impulsivity and cigarette smoking: delay discounting in current, never, and ex-smokers. *Psychopharmacology, 146*(4), 447-454.

Bock, R. D., & Aitkin, M. (1981). Marginal maximum-likelihood estimation of item parameters - Application of an EM algorithm. *Psychometrika, 46*(4), 443-459.

Chao, L.-W., Szrek, H., Pereira, N. S., & Pauly, M. V. (2009). Time preference and its relationship with age, health, and survival probability. *Judgment and Decision Making, 4*(1), 1-19.

Chapman, G. B. (1996). Temporal discounting and utility for health and money. *Journal of Experimental Psychology-Learning Memory and Cognition, 22*(3), 771-791.

Chapman, G. B., & Elstein, A. S. (1995). Valuing the future - temporal discounting of health and money. *Medical Decision Making, 15*(4), 373-386.

Chen, S. K., Hou, L. L., & Dodd, B. G. (1998). A comparison of maximum likelihood estimation and expected a posteriori estimation in CAT using the partial credit model. *Educational and Psychological Measurement, 58*(4), 569-595.

Cormen, T. H., Leiserson, C. E., & Rivest, R. L. (1990). *Introduction to Algorithms* (1 ed.): MIT Press and McGraw-Hill.

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*(2), 202-207.

De Ayala, R. J. (1999). Item parameter recovery for the nominal response model. *Applied Psychological Measurement, 23*(1), 3-19.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction with Life Scale. *Journal of Personality Assessment, 49*(1), 71-75.

Doyle, J. R. (2013). Survey of time preference, delay discounting models. *Judgment and Decision Making, 8*(2), 116-135.

Forero, C. G., & Maydeu-Olivares, A. (2009). Estimation of IRT Graded Response Models: Limited Versus Full Information Methods. *Psychological Methods, 14*(3), 275-299.

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature, 40*(2), 351-401.

Friedel, J. E., DeHart, W. B., Madden, G. J., & Odum, A. L. (2014). Impulsivity and cigarette smoking: discounting of monetary and consumable outcomes in current and non-smokers. *Psychopharmacology, 231*(23), 4517-4526.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., et al. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality, 40*(1), 84-96.

Green, L., Fristoe, N., & Myerson, J. (1994). Temporal discounting and preference reversals in choice between delayed outcomes. *Psychonomic Bulletin & Review, 1*(3), 383-389.

Green, L., Fry, A. F., & Myerson, J. (1994). Discounting of delayed rewards: A life-span comparison. *Psychological Science, 5*(1), 33-36.

Griskevicius, V., Tybur, J. M., Delton, A. W., & Robertson, T. E. (2011). The Influence of Mortality and Socioeconomic Status on Risk and Delayed Rewards: A Life History Theory Approach. *Journal of Personality and Social Psychology, 100*(6), 1015-1026.

Hirsh, J. B., Morisano, D., & Peterson, J. B. (2008). Delay discounting: Interactions between personality and cognitive ability. *Journal of Research in Personality, 42*(6), 1646-1650.

Horvath, A. (2012). Binary search and linear search performance on the .NET and Mono platform. from http://blog.teamleadnet.com/2012/02/quicksort-binary-search-and-linear.html

Johnson, M. W. (2012). An Efficient Operant Choice Procedure for Assessing Delay Discounting in Humans: Initial Validation in Cocaine-Dependent and Control Individuals. *Experimental and Clinical Psychopharmacology, 20*(3), 191-204.

Johnson, M. W., & Bickel, W. K. (2002). Within-subject comparison of real and hypothetical money rewards in delay discounting. *Journal of the Experimental Analysis of Behavior, 77*(2), 129-146.

Kirby, K. N. (1997). Bidding on the future: Evidence against normative discounting of delayed rewards. *Journal of Experimental Psychology-General, 126*(1), 54-70.

Kirby, K. N., & Marakovic, N. N. (1996). Delay-discounting probabilistic rewards: Rates decrease as amounts increase. *Psychonomic Bulletin & Review, 3*(1), 100-104.

Kirby, K. N., & Petry, N. M. (2004). Heroin and cocaine abusers have higher discount rates for delayed rewards than alcoholics or non-drug-using controls. *Addiction, 99*(4), 461-471.

Kirby, K. N., Winston, G. C., & Santiesteban, M. (2005). Impatience and grades: Delay-discount rates correlate negatively with college GPA. *Learning and Individual Differences, 51*, 563-569.

Krishnan-Sarin, S., Reynolds, B., Duhig, A. M., Smith, A., Liss, T., McFetridge, A., et al. (2007). Behavioral impulsivity predicts treatment outcome in a smoking cessation program for adolescent smokers. *Drug and Alcohol Dependence, 88*(1), 79-82.

Lagorio, C. H., & Madden, G. J. (2005). Delay discounting of real and hypothetical rewards III: Steady-state assessments, forced-choice trials, and all real rewards. *Behavioural Processes, 69*(2), 173-187.

Lane, S. D., Cherek, D. R., Pietras, C. J., & Tcheremissine, O. V. (2003). Measurement of delay discounting using trial-by-trial consequences *Behavioural Processes*, *64*, 287-303.

Lawyer, S. R., Schoepflin, F., Green, R., & Jenks, C. (2011). Discounting of Hypothetical and Potentially Real Outcomes in Nicotine-Dependent and Nondependent Samples. *Experimental and Clinical Psychopharmacology, 19*(4), 263-274.

Locey, M. L., Jones, B. A., & Rachlin, H. (2011). Real and hypothetical rewards in self-control and social discounting. *Judgment and Decision Making, 6*(6), 552-564.

Madden, G. J., Begotka, A. M., Raiff, B. R., & Kastern, L. L. (2003). Delay discounting of real and hypothetical rewards. *Experimental and Clinical Psychopharmacology, 11*(2), 139-145.

Madden, G. J., Raiff, B. R., Lagorio, C. H., Begotka, A. M., Mueller, A. M., Hehli, D. J., et al. (2004). Delay discounting of potentially real and hypothetical rewards: II. Between- and within-subject comparisons. *Experimental and Clinical Psychopharmacology*, *12*(4), 251-261.

Magen, E., Dweck, C. S., & Gross, J. J. (2008). The hidden-zero effect - Representing a single choice as an extended sequence reduces impulsive choice. *Psychological Science*, *19*(7), 648-649.

Magis, D., & Mahalingam, V. (2015). Computer adaptive testing. In U. o. Liege & U. o. Cambridge (Eds.), *Applicações de Métodos Estatísticos Avançadaos à Avaliação Psicológica e Educacional* (pp. 239-256). São Paulo: Vetor Editoria.

Mahalingam, V., Stillwell, D., Kosinski, M., Rust, J., & Kogan, A. (2014). Who can wait for the future? A personality perspective. *Social Psychological and Personality Science, 5*(5), 573 - 583.

Matusiewicz, A. K., Carter, A. E., Landes, R. D., & Yi, R. (2013). Statistical equivalence and test-retest reliability of delay and probability discounting using real and hypothetical rewards. *Behavioural Processes, 100*, 116-122.

Michaelson, L., de la Vega, A., Chatham, C. H., & Munakata, Y. (2013). Delaying gratification depends on social trust. *Frontiers in Psychology, 4*.

Myerson, J., Green, L., & Warusawitharana, M. (2001). Area under the curve as a measure of discounting. *Journal of the Experimental Analysis of Behavior, 76*(2), 235-243.

Odum, A. L. (2011). Delay discounting: I'm a k, you're a k. *Journal of the Experimental Analysis of Behavior, 96*(3), 427-439.

Ostaszewski, P. (1996). The relation between temperament and rate of temporal discounting. *European Journal of Personality, 10*(3), 161-172.

Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt Impulsiveness Scale. *Journal of Clinical Psychology, 51*(6), 768-774.

Rachlin, H., Raineri, A., & Cross, D. (1991). Subjective-probability and delay. *Journal of the Experimental Analysis of Behavior, 55*(2), 233-244.

Raineri, A., & Rachlin, H. (1993). The effect of temporal constraints on the value of money and other commodities. *Journal of Behavioral Decision Making, 6*(2), 77-94.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods* (2 ed.). Chicago: SAGE.

Reimers, S., Maylor, E. A., Stewart, N., & Chater, N. (2009). Associations between a one-shot delay discounting measure and age, income, education and real-world impulsive behavior. *Personality and Individual Differences, 47*(8), 973-978.

Reynolds, B., Karraker, K., Horn, K., & Richards, J. B. (2003). Delay and probability discounting as related to different stages of adolescent smoking and non-smoking. *Behavioural Processes, 64*(3), 333-344.

Reynolds, B., Ortengren, A., Richards, J. B., & de Wit, H. (2006). Dimensions of impulsive behavior: Personality and behavioral measures. *Personality and Individual Differences, 40*(2), 305-315.

Reynolds, B., Patak, M., Shroff, P., Penfold, R. B., Melanko, S., & Duhig, A. M. (2007). Laboratory and self-report assessments of impulsive behavior in adolescent daily smokers and nonsmokers. *Experimental and Clinical Psychopharmacology, 15*(3), 264-271.

Reynolds, B., Richards, J. B., Horn, K., & Karraker, K. (2004). Delay discounting and probability discounting as related to cigarette smoking status in adults. *Behavioural Processes, 65*(1), 35-42.

Richards, J. B., Zhang, L., Mitchell, S. H., & de Wit, H. (1999). Delay or probability discounting in a model of impulsive behavior: Effect of alcohol. *Journal of the Experimental Analysis of Behavior, 71*(2), 121-143.

Rotter, J. B. (1967). A new scale for the measurement of interpersonal trust. *Journal of personality, 35*(4), 651-665.

Samuelson, P. (1937). A note on measurement of utility. *The Review of Economic Studies, 4*(2), 155-161.

Sircova, A., & Mitina, O. (2008). Zimbardo time perspective inventory : A comparison across cultures. *International Journal of Psychology, 43*(3-4), 43-43.

Smith, V. K., Taylor, D. H., & Sloan, F. A. (2001). Longevity expectations and death: Can people predict their own demise? *American Economic Review, 91*(4), 1126-1134.

Stanford, M. S., Mathias, C. W., Dougherty, D. M., Lake, S. L., Anderson, N. E., & Patton, J. H. (2009). Fifty years of the Barratt Impulsiveness Scale: An update and review. *Personality and Individual Differences, 47*(5), 385-395.

Stillwell, D. J., & Kosinski, M. (2011). myPersonality Research Wiki. from http://mypersonality.org/wiki

Stillwell, D. J., & Tunney, R. J. (2012). Effects of measurement methods on the relationship between smoking and delay reward discounting. *Addiction, 107*(5), 1003-1012.

Stolarski, M., Bitner, J., & Zimbardo, P. G. (2011). Time perspective, emotional intelligence and discounting of delayed awards. *Time & Society, 20*(3), 346-363.

Takahashi, T., Ikeda, K., & Hasegawa, T. (2007). A hyperbolic decay of subjective probability of obtaining delayed rewards. *Behavioral and Brain Functions, 3*.

Wainer, H. (2010). Fourteen conversations about three things. *Journal of Educational and Behavioral Statistics, 35*, 5-25.

Wang, T. Y., & Vispoel, W. P. (1998). Properties of ability estimation methods in computerized adaptive testing. *Journal of Educational Measurement, 35*(2), 109-135.

Wang, W. C., & Chen, C. T. (2005). Item parameter recovery, standard error estimates, and fit statistics of the Winsteps program for the family of Rasch models. *Educational and Psychological Measurement, 65*(3), 376-404.

Weller, R. E., Cook, E. W., III, Avsar, K. B., & Cox, J. E. (2008). Obese women show greater delay discounting than healthy-weight women. *Appetite, 51*(3), 563-569.