

Other Papers

Polish Psychological Bulletin
2017, vol. 48(1) 118–128
DOI - 10.1515/ppb-2017-0015

Marta Malesza*

Testing the construct validity of the Discounting Inventory – Psychometric properties of a Polish and German samples

Abstract: *The Discounting Inventory (DI), originally developed in Polish language, allows the measurement of individual differences in the delay, probabilistic, effort, and social discounting rates. The present study attempted to validate the DI's psychometric properties using German university students and to compare the results to those from a sample of Polish university students. Over four hundred participants completed the DI and traditional discounting measures. A confirmatory factor analysis indicated that the original four-factor model of the DI provided an excellent fit for the German data and internal consistency was high. These outcomes were similar, if not superior, to those from the Polish sample. DI scores strongly correlated with traditional discounting measures scores in both samples, replicating previous results. These findings indicate that the DI is a valid measure for use in a sample from another cultural setting, which is potentially useful to both researchers and practitioners.*

Key words: *Discounting Inventory, construct validity, behavioral impulsivity, university students*

Introduction

Impulsivity – i.e. the tendency to act with little forethought, without deliberation and evaluation of consequences – is considered as a major trait of personality. Given the fact that impulsivity is multidimensional in nature (Ainslie, 1975; Kirby & Finch, 2010; Reynolds, Ortengren, Richards, & de Wit, 2006), a growing number of researchers have used a behavioral definition of impulsivity, according to which impulsivity is a tendency to choose smaller, more immediate rewards over larger, more delayed rewards (Ainslie, 1975; Rachlin, Raineri, & Cross, 1991). The value of the bigger reward is said to have been discounted. The term “discounting process” refers to a decrease in the subjective value of an outcome as a specific environmental factor on which a reward or a loss is devalued increases (e.g., Green & Myerson, 2004; Rachlin et al., 1991). The most widely studied process, delay discounting (also known as temporal discounting; for review, see Madden & Bickel, 2010) refers to the

behavioral definition of impulsivity mentioned above – the preference for smaller outcomes that are available relatively sooner over larger outcomes that are available after a delay (Ainslie, 1975). Of course, the value of a reward decreases as a function of variables other than time (see: Green & Myerson, 1996; Myerson & Green, 1995; Ostaszewski, 1997; Rachlin et al., 1991). Apart from the discounting of delayed rewards, behavioral psychology also studies probabilistic discounting (the process by which the subjective value of the gain diminishes together with the decreasing probability of achieving the gain; Ostaszewski, Green, & Myerson, 1998), and effort discounting (the decrease in subjective value of the gain coinciding with the increasing effort needed for gaining the reward; Mitchell, 2004; Sugiwaka & Okouchi, 2004), as well as social discounting (defined as the process by which the subjective value of the reinforcement diminishes according to the increasing number of people with whom the reward is to be shared; Rachlin, 1993).

* University of Warsaw, Faculty of Psychology

Corresponding author: marta.malesza@psych.uw.edu.pl

Funding source with grant number

This research was supported by a grant from the National Science Center in Poland (grant number: 2013/11/N/HS6/01149; grant: “Discounting as a personality trait: Construction and validation of the Discounting Inventory, with special focus on the research concerning DI’s diagnostic validation”; Head of the project: Marta Malesza).

Measurement methods of discounting

There have been numerous attempts to develop screening measures to identify the potential presence of steep discounting rate (e.g., Navarick, 2004; Smith & Hantula, 2008). To evaluate the discounting of delayed rewards, the most commonly used traditional discounting measure presents an individual with a series of pairs of hypothetical choices: participants choose between a smaller, more immediate alternative and a larger, more delayed alternative (e.g., Green & Myerson, 2004; Rachlin et al., 1991). For example, suppose a participant is presented with a choice between receiving \$10 immediately or \$100 in six months, and he chooses the immediate option, and that subsequently the participant must decide between \$10 or \$110 in six months, and he chooses the future option (Rachlin et al., 1991). This pattern of choices implies that the participant would be indifferent between \$10 today and roughly \$105 in six months. This indifference point can then be converted into a discount rate using a number of different models (e.g., Myerson, Green, & Warusawitharana, 2001). Second, the matching method, in contrast to the choice-based method, asks for the indifference point directly. For example, it might ask the participant what amount “X” would make him indifferent between \$10 immediately and \$X in six months.

Another common measure to assess delay discounting is the Monetary Choice Questionnaire (MCQ) developed by Kirby, Petry, and Bickel (1999). Participants are asked to make 27 choices between smaller immediate rewards versus larger delayed rewards, and the degree of discounting is interpolated from their choices (Kirby et al., 1999). In this procedure that uses hypothetical outcomes, indifference points are determined by adjusting the amount of the immediate outcome based on the participant’s choice, rather than by moving through a fixed list of options as in Rachlin et al. (1991). While the MCQ has great utility because of its simplicity and brief administration time. However, it is possible that the MCQ produces a ceiling effect in estimating delay discounting parameters in highly impulsive individuals.

The hypothetical money choice tasks described in the paragraphs above have generated skepticism and attempts to develop other procedures with better face validity. In the Experiential Discounting Task (EDT; Reynolds & Schiffbauer, 2004), for example, participants experience the delays and amounts of money that they choose. On each trial, participants decide between a delayed and uncertain standard amount of money (\$ 0.30) and an immediate adjusting amount. Across blocks of trials, the delay to the standard amount is changed to determine a discount function. The EDT differs from other delay discounting procedures by making the standard reward probabilistic in addition to delayed, and thus examines the simultaneous effects of delay and certainty on reward value. However, amounts of money participants receive are rather small and the delays participants experience are rather short. Thus, it can be doubt it these features of the EDT make it correspond more closely to situations that people experience in their lives outside the laboratory (Reynolds, 2006).

As has been seen, typical discounting measurement methods are somehow different from the usual forms of psychometric assessment. Although widely used, some researchers have argued that the traditional discounting measure suffers from a number of practical problems (Navarick, 2004). Obviously, every researcher who has used hypothetical rewards has questioned the validity of their procedures, noting that choices made between these outcomes may not accurately reflect the choices between real outcomes (Madden et al., 2004). The additional problem with traditional discounting measures using pairs of hypothetical choices is that the accuracy of measurement may be compromised due to task fatigue or boredom as a result of the many choices required, e.g., 100 or more (Navarick, 2004). For these and other reasons, resent research was devoted to constructing a tool different from traditional means of measuring the discounting rate consisting of pairs of hypothetical choices (Malesza & Ostaszewski, in press).

Discounting Inventory

Malesza and Ostaszewski (in press) introduced a Discounting Inventory that allows the measurement of individual differences in the delay, probabilistic, effort, and social discounting rates. The construction of the Discounting Inventory (DI) comprised a variety of steps. The starting point of the research consisted of a thorough theoretical analysis of all concepts that refer to the discounting process (for details, see Murphy & Davidshofer, 2005). Over 400 items covering four types of discounting were generated. Next, a thorough psychometric study and factor analysis of data obtained from a group of 2843 subjects allowed them to test the DI’s construct validity. These 2843 respondents were divided into two groups and an Exploratory Factor Analyses was conducted on the data from the first group, and Confirmatory Factor Analyses was conducted on the data from the second group. Results from the Exploratory Factor Analyses indicated a four-factor solution. Confirmatory Factor Analyses, using structural equation modeling (Maximum Likelihood), was used to confirm the factor structure of the data, and these analyses indicated that the four-factor structure proposed had the best fit to the data. These factors were closely associated with the theoretical four dimensions, which we have referred to as delay discounting, probabilistic discounting, social discounting, and effort discounting (Green & Myerson, 2004; Ostaszewski et al., 1998; Rachlin, 1993; Sugiwaka & Okouchi, 2004). Additionally, to meet the need for a shorter instrument that assesses all four types of discounting efficiently, the authors decided to reduce the remaining pull of items. Through several iterations of retaining and deleting items on the basis of their component loadings, item intercorrelations, and contribution to coefficient alphas, the total number of items was reduced from 209 to 48 (12 items per scale; for details see Malesza & Ostaszewski, in press).

Previous psychometric work has demonstrated that the internal consistency, temporal reliability, and construct validity of the DI are all adequate (Malesza & Ostaszewski, in

press). The most fundamental of these properties is reliability, which establishes the upper bound for validity (Anastasi & Urbina, 1997). Two types of reliability were of interest. So far, the test–retest reliability of the DI measure was assessed during a 2-week period. Despite being collected two weeks apart, a recent analysis indicated that each of the participants reports of the four dimensions exhibited moderate to strong test stability ($r_{tt} = 0.65\text{--}0.82$). All reliabilities were significant with $p < 0.05$. Hence, Malesza and Ostaszewski (in press) concluded that the test–retest reliability results suggested that items intended to measure four types of discounting were likely testing trait, rather than state, factors. Second, the internal consistencies of the DI measured with Cronbach's alphas (Cronbach, 1990) are also adequate and relatively high (0.82–0.95). In case of validity, it was also important to evaluate if the DI measures the same construct as traditional discounting instruments. Significant correlations between DI and traditional discounting measures were reported (Malesza & Ostaszewski, in press: $r = 0.20\text{--}0.47$). The final 48-item version of the inventory has satisfactory psychometric characteristics, and seems to be a relevant alternative to the traditional discounting measures.

Cross-cultural validation of research measures

Although the DI was designed to measuring discounting, one cannot assume that instruments designed and tested in one cultural setting will necessarily function as intended in a different culture (Hambleton, 2001). That is, prior research has demonstrated that the DI identifies discounting behaviors in a Polish sample (Malesza & Ostaszewski, in press), but such findings do not guarantee that the psychometric characteristics will be retained when the DI is used in a sample from another cultural setting. Likewise, although previous study has demonstrated that endorsing discounting measured by the DI is significantly related to the traditional discounting measure of discounting (Malesza & Ostaszewski, in press) in Polish participant, one cannot assume that such a relationship will be universally found in different cultures, because people may differ in the reasons for why they show steep discounting (i.e. are more impulsive). With that said, if a particular measure retains sound psychometric characteristics when used in different cultures, then that measure has several things to recommend it (Hambleton, 2001; Murphy & Davidshofer, 2005). First, it would provide a single measure that was potentially useful to researchers and practitioners in multiple cultures. Second, such a research should confirm its psychometric characteristics when used outside of the culture in which it was originally developed.

Overview of the present research

To that end, the present study recruited German students attending two universities in Germany to complete the DI. As a means of comparison, data were also collected from students attending one university in Poland. The major purpose of the study was to determine whether the psychometric structure identified by Malesza and Ostaszewski (in press) would describe the data from both the Polish and German samples. In this research,

the author was especially interested in two psychometric characteristics. The first was the factor structure of the individual items in the DI. It was predicted that, as in Malesza and Ostaszewski (in press), 48 items intended to measure four types of discounting would load strongly onto four factors. The second characteristic was the internal consistency. It was predicted that the internal consistency measures would be good (> 0.80) and comparable to those reported by Malesza and Ostaszewski (in press).

A secondary goal of the present study was to determine whether the significant relationship between DI and traditional discounting measures using pairs of hypothetical choices would be replicated in both the Polish and German samples. To accomplish that goal, all of the participants in the German and Polish samples also completed the traditional discounting measures (Richards, Zhang, Mitchell, & de Wit, 1999). Both DI and traditional discounting measures are intended to assess the discounting construct, so the two measures should be correlated. However, they differ in the way they assess the construct of discounting, including self-reported measures of personality that rely on individuals' perceptions of their behavior and behavioral tasks that measure overt behavior related to specific dimensions of impulsivity (Reynolds et al., 2006). In addition, previous studies have shown modest significant correlations between different measures of discounting (e.g., DI and traditional measure of the discounting rate: $r = 0.20\text{--}0.47$; Experiential Discounting Task and traditional measure of the discounting rate: $r = 0.26\text{--}0.52$; Malesza & Ostaszewski, in press; Reynolds, 2006; Reynolds & Schiffbauer, 2004). Thus, the hypothesis was that scores on the DI scales would be significantly correlated with scores on the traditional discounting instruments measuring four types of discounting. Correlations were expected to be in the small to moderate range ($r = 0.30\text{--}0.50$).

Participants

The total sample consisted of 482 participants. The Polish participants were 208 (145 female; 63 male) students. The mean age of these students was 23.5 years ($SD = 1.8$ years; range = 18–36 years). All participants volunteered their time for free for their participation in the study. The German participants were 274 (200 female; 74 male) students. The mean age of these students was 22.1 years ($SD = 3.2$ years; range = 18–27 years). All participants received extra course credit in their psychology course in return for their participation.

All participants gave their informed consent before inclusion in the study. The local Institutional Review Board approved the study, and participants were treated according to the Ethical Principles of Psychologists and Code of Conduct (American Psychiatric Association, 1992). The Polish and German participants experienced the exact same procedure, with one exception. All Polish students completed the paper-and-pen materials, while data from German participants were collected in an online study administered via a tool for online surveys: www.soscisurvey.de.

Materials and Procedure

Discounting Inventory

The first item completed by all participants was the DI (Malesza & Ostaszewski, in press). The measurement comprised the 48-item pool with the 4-point Likert-like scale format (4 = fully agree, 3 = agree slightly, 2 = disagree slightly, 1 = disagree completely). Responses for four 12-item scales are summed (i.e., no items are reverse coded) to provide a score for that particular scale.

The Discounting Inventory was originally developed in Polish language. To ensure that the items resemble the meaning of the original German items¹ as closely as possible, we followed a common procedure of back-translation in which a text is translated from a source into a target language, and a second interpreter independently translates the text back into the source language. The Polish version of the measure was first translated into German and then back-translated into Polish by two translators, according to the guidelines developed by the International Test Commission (Hambleton, 2001). Afterwards, the accuracy of the translation was evaluated by comparing the original and back-translated versions.

Traditional discounting measure

Participants also completed the traditional discounting measure using pairs of hypothetical choices adapted from Richards et al. (1999). Measure consisted of four parts. One assessed the steepness of delay discounting, the second one assessed the steepness of effort discounting, the third one the steepness of social discounting, and the fourth one assessed the steepness of probabilistic discounting. The sequence of tasks was counterbalanced across participants.

The effort discounting part consisted of five pages, each with a table with two columns. Effortful reward was presented in the right-hand column, together with information about particular effort requirements. Effortless rewards were presented in the left-hand column. On each page, a different value of effort necessary to receive a reward of PLN 800 (for German participants all sums were presented in euro currency; at the time of the study, €1 ≈ PLN 4.00) was presented. On consecutive pages, the values of effort increased. The extent of the effort depended on the floor to which the participant had to climb. Participants were asked to imagine climbing stairs up to a specified floor (the 3rd, 10th, 15th, 30th, and 50th floor). The effortless rewards were printed in rows in descending order, starting at 100% of the value of an effortful reward and ending at 0% of the effortful reward. For example, in the fifth row of one table, a participant had to choose between climbing to the 30th floor to receive PLN 800 or an effortless reward of PLN 680. Participants were asked to mark which of the two rewards (effortful or effortless) they chose in every row of every table.

The delay, probabilistic, and social discounting parts of the traditional discounting questionnaire were prepared

in exactly the same way as the effort discounting part. The delayed reward (PLN 1,400) was presented in the right-hand column, together with the information about the particular delay (1 mo., 6 mo., 12 mo., 5 yr., or 15 yr.). Immediate rewards were presented in the left-hand column and printed in rows in descending order from 1,400 to 0 PLN. Again, participants were asked to mark which of the two rewards (delayed or immediate) they chose in every row of every table. Probabilistic discounting – in this case, the probability of receiving a reward – was assessed at five probability interval values: 5%, 25%, 50%, 75%, and 95%. On each probabilistic task trial, participants chose between a certain amount of money and the possibility of receiving PLN 1,200 with a specified probability. In the social discounting portion, participants made hypothetical choices between a smaller monetary reward exclusively for themselves or a larger reward they had to share equally with a specific number of strangers (1, 3, 5, 11, 17 people; PLN 900).

Although the delay, probabilistic, social, and effort tasks and the outcomes were hypothetical, participants were instructed to act as if the situation were real. Participants were told: *You will not receive any of the rewards that you choose, but we want you to make your decisions as though you were really going to get the rewards you choose.*

Procedure

Participants started each page of the traditional discounting measure from the top, where both amounts were equal, and chose one of the two options in each row. The aim of the procedure was to discover the lowest amount of effortless, certain, received for oneself, and immediate reward that a participant would prefer to receive instead of a reward that required a particular effort, probability, sharing with others, or delay. This lowest amount would be the last amount that the participant chose in the left-hand column, before switching to the effortful, uncertain, shared, or delayed option in the right-hand column. This amount was considered the subjective value of the reward for a given magnitude of effort, probability, number of people, or given delay.

Analysis of the data from the traditional discounting measure

An area-under-the-curve (AUC) method was used to characterize the delay, probabilistic, social, and effort discounting rate (Myerson et al., 2001). AUC involves computing the area of the trapezoids that are created by plotting the coordinates of indifference points for each delay period, probability, number of people, and effort values. The equation for the area of the trapezoids is $(x_2 - x_1) * [(y_2 + y_1) / 2]$, where the ordinate represents the proportion of absolute (non-discounted) reward value, and the abscissa represents the proportion of maximum possible delay to reinforcement, probability of getting the reward, number of people sharing the reward, or effort required to get the reinforcement (Myerson et al., 2001). These values

¹ German version of the Discounting Inventory can be found in the Appendix.

are summed to obtain a total area that ranges from 0.0 to 1.0; larger AUC values are indicative of slower or no discounting, and lower AUC values mean greater levels of discounting.

Myerson et al. (2001) argued that AUC has several merits that make it appropriate for use in discounting research. One advantage is that the AUC measure is theoretically neutral and can be calculated for all participants regardless of their response pattern. AUC requires no a priori assumptions about the shape of the discount function or the number of free parameters used in modeling. In addition, AUC measure usually follows a normal distribution, which permits the use of parametric statistical analysis. As such, these improved measurement properties make it a stronger candidate marker than discounting rates.

Individual hypothetical discounting patterns were also categorized as systematic and nonsystematic on the basis of Johnson and Bickel's (2008) criteria of identifying atypical response patterns that suggest random or inconsistent patterns of responding in both Polish and German samples. Specifically, individual participants were considered nonsystematic responders if the analysis of their hypothetical discounting data revealed that (1) any indifference point (except for the first one) was larger than the previous one by more than 10% and (2) the last indifference point was not less than the first by at least 10%.

In the Polish sample, 19 (9%) data sets out of the 208 totally examined was identified as nonsystematic due to criterion 1. In each of these data sets, at least one indifference point suggests a departure from the monotonically decreasing function. Four data sets (2%) out of 208 examined were identified as nonsystematic due to criterion 2. That is, the last indifference point was not less than the first indifference point by at least a magnitude equal to 10% of the larger later reward. Inspection of data sets suggests that these participants were not sensitive to the effort condition. In the German sample, 22 (8%) and seven (2,5%) data sets out of these 274 totally examined were identified as nonsystematic due to criterion 1 and 2, respectively.

Results

Internal consistency

The internal consistencies measured with Cronbach's alphas are adequate. The following coefficients for the internal consistency for the Polish group of participants were observed: total measure $\alpha = 0.92$, effort discounting scale $\alpha = 0.86$, probabilistic discounting scale $\alpha = 0.90$, social discounting scale $\alpha = 0.91$, and delay discounting scale $\alpha = 0.89$, suggesting that the items have relatively high internal consistency. For comparison, when the data from all 274 German participants were analysed, the Cronbach's alphas for the German students were 0.96, 0.93, 0.87, 0.84, and 0.92 for the overall DI, delay, effort, probabilistic, and social discounting scales, respectively.

DI Factor Structure

Next, we evaluated if the previously reported four-factor structure of the DI was valid for our data. The data generated by Polish and German respondents were subjected to confirmatory factor analyses, using the AMOS program (Arbuckle, 1997). The estimation method was Maximum Likelihood (ML). Note that, ML makes assumption about multivariate normality (Lei & Lomax, 2005). As both skewness and kurtosis were less than or equal to 1.2 for all items, indicating that the item-distributions were similar to the rest of the items in the instrument and that the item distributions were rather symmetrical, the ML estimation was considered appropriate.

In these analyses, two models were examined: (1) one-factor solution (all 48 items were combined into one factor); and (2) four-factor solution, which comprised four factors representing four types of discounting. The choice of models was influenced by Malesza and Ostaszewski (in press). Specifically, these two models above examined the factor structure of the original DI (see Malesza & Ostaszewski, in press).

Model fit

Model fit was assessed using multiple indices (see Mulaik, 2007). A well-fitting model should ideally have a nonsignificant χ^2 statistic ($p > 0.05$). Because the χ^2 statistic tends to be inflated in large samples (> 200 subjects), the ratio χ^2/df was determined, which should not be much larger than 2.0. The χ^2/df is a measure of the absolute fit of the model with the data, indicating how closely the model fits compared to a perfect fit. The model is considered to have an excellent fit if the GFI (goodness-of-fit index), TLI (Tucker–Lewis index), and CFI (comparative fit index) values are approximately 0.95 or above (or 0.90–0.95 for an acceptable fit). The RMSEA (root mean square error of approximation) represents reasonable errors of approximation in the population; a value of approximately 0.05 or less would indicate a close fit, and a value of up to 0.08 would represent a reasonable fit of the model. Note that, however, the choice of indices and cutoff values is a topic surrounded by considerable controversy (see, e.g., Mulaik, 2007).

Factor structure – Polish sample

The four-factor model had a good fit to the data from Polish sample: χ^2 (25 df) = 50.82, $p = 0.03$; $\chi^2/df = 2.03$; RMSEA = 0.03, GFI = 0.98, CFI = 0.99, TLI = 0.96. Although the Chi square test of model fit was significant, the other fit indices indicated an excellent model fit. Second, for the purpose of comparison, a one-factor model, which presupposes that all the items pertain to the same factor, was also assessed. According to the χ^2 statistic, the model would have to be rejected (χ^2 [69 df] = 217.0, $p < .001$). The value $\chi^2/df = 3.14$ obtained here also indicates an unacceptable fit. An acceptable χ^2/df is usually set at or less than 2. The four-factor model had a much better fit than a one-factor model of general discounting, $\Delta\chi^2_{(44)} = 166.18, p < .001$.

Factor structure – German sample

Next, the data from the remaining 274 German participants were subjected to a Confirmatory Factor Analyses identical to that used for the Polish sample. The fit indices were as follows: χ^2 (106 df) = 214.58, $p = 0.01$; $\chi^2/df = 2.02$; RMSEA = 0.07; GFI = 0.99, CFI = 0.97, TLI = 0.99. The outcome of the CFA, suggested that a four-factor solution yielded the best model fit. As with the data from Polish sample, CFI, GFI, TLI, RMSEA, and χ^2/df , respectively, indicated a good or an acceptable model fit, whereas according to the χ^2 statistic, the model would have to be rejected. Finally, the χ^2 difference test yielded a better fit for four-factor model than one factor model ($\Delta\chi^2_{(182)} = 538, p < .001$).

Also, in both samples, based on the standardized regression weights, each item was linked to a single factor. All of these coefficients exceed 0.4, providing additional support for the efficacy of this model. The standardized coefficient associated with each item is given in Table 2.

Correlations with traditional discounting measures

To verify the relationship between two measures of discounting, Pearson's correlation coefficients were calculated. The correlations were also run for each group separately. Table 3 summarizes these results. All correlations are in the expected direction. Across groups, the correlation between discounting scores in the Polish sample varied from -0.19 to -0.48, while in the German sample varied from -0.23 to -0.51.

DI produced significant negative correlations with the established measures of discounting. In the more

easily interpreted scenario, the results showed that the higher one's scores on the DI's scales, the steeper one's delay, probabilistic, social, and effort discounting were. In addition, all four DI scales provided significant associations with each discounting type measured by the traditional discounting counterpart. For example, in the German sample, the convergent validity correlation for the DI delay scale with the standard measure of delay discounting ($r = -0.51$) was generally larger than its correlations with three other traditional discounting scales (effort $r = -0.36$, probability $r = -0.39$, social $r = -0.29$). However, the DI probability discounting scale manifested a significant correlation with the established measure of delay discounting ($r = -0.46$) and was generally larger than its correlations with well-established traditional probabilistic ($r = -0.35$) discounting measure. By contrast, in the Polish sample, all DI scales produced higher significant correlations with their standard discounting counterpart than its correlations with three other traditional discounting types.

General discussion

The present study evaluated the construct validity of the Discounting Inventory. Construct validity assesses whether a test actually measures the construct it purports to measure (Cronbach, 1990), and is shown by within-scale analyses and analyses against external criteria. Within-scale analyses should show evidence that a single entity (construct) is being measured and that items can be combined to form a summary score (Murphy &

Table 1. Confirmatory Factor Analysis fit

	Polish sample							
	χ^2	df	χ^2/df	p	RMSEA	GFI	CFI	TLI
Model 0	308.1	69	4.46	0.001	0.06	0.83	0.90	0.85
Model 1	217.0	69	3.14	0.001	0.06	0.90	0.92	0.89
Model 2	179.5	60	2.99	0.010	0.05	0.91	0.94	0.91
Model 3	135.6	58	2.34	0.010	0.05	0.91	0.93	0.94
Model 4	50.82	25	2.03	0.030	0.03	0.98	0.99	0.96
	German sample							
	χ^2	df	χ^2/df	p	RMSEA	GFI	CFI	TLI
Model 0	802.30	360	2.23	0.001	0.08	0.85	0.89	0.80
Model 1	752.58	288	2.61	0.001	0.06	0.89	0.90	0.84
Model 2	590.50	204	2.89	0.001	0.06	0.90	0.91	0.88
Model 3	432.71	197	2.20	0.010	0.07	0.87	0.90	0.89
Model 4	214.58	106	2.02	0.010	0.07	0.99	0.97	0.99

Note: Model 0 = null-factor solution; Model 1 = one-factor solution; Model 2 = two-factor solution; Model 3 = three-factor solution; Model 4 = four-factor solution; GFI = goodness-of-fit index; TLI = Tucker–Lewis index; CFI = comparative fit index; RMSEA = root mean square error of approximation.

Table 2. Standardized loadings for the four-factor confirmatory models of the Discounting Inventory items

No. of item	Polish sample				German sample			
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 2	Factor 3	Factor 4
1.	0.76				0.66			
5.	0.63				0.76			
9.	0.70				0.55			
13.	0.59				0.62			
17.	0.72				0.60			
21.	0.57				0.64			
25.	0.81				0.71			
29.	0.92				0.81			
33.	0.83				0.66			
37.	0.55				0.69			
41.	0.50				0.53			
45.	0.43				0.88			
2.		0.69				0.49		
6.		0.85				0.45		
10.		0.91				0.88		
14.		0.46				0.67		
18.		0.83				0.63		
22.		0.74				0.59		
26.		0.49				0.88		
30.		0.51				0.55		
34.		0.46				0.44		
38.		0.69				0.91		
42.		0.67				0.90		
46.		0.89				0.85		
3.			0.70				0.75	
7.			0.48				0.81	
11.			0.72				0.69	
15.			0.91				0.52	
19.			0.87				0.45	
23.			0.80				0.44	
27.			0.85				0.49	
31.			0.83				0.57	
35.			0.74				0.59	
39.			0.76				0.48	
43.			0.79				0.51	
47.			0.51				0.76	
4.				0.68				0.65
8.				0.73				0.63
12.				0.64				0.68
16.				0.56				0.49
20.				0.59				0.41
24.				0.62				0.52
28.				0.50				0.57
32.				0.71				0.60
36.				0.95				0.85
40.				0.57				0.69
44.				0.88				0.80
48.				0.61				0.91

Note: Factor 1 = Delay Discounting scale; Factor 2 = Probabilistic Discounting scale; Factor 3 = Effort Discounting scale; Factor 4 = Social Discounting scale that emerged from the Confirmatory Factor Analyses.

Table 3. Correlation matrix comparing all measures of discounting

Polish sample								
Discounting Inventory measure				Traditional discounting measure				
	1	2	3	4	5	6	7	8
1	–	0.49**	0.39**	0.20*	-0.44**	-0.31**	-.27*	-.29*
2		–	0.28*	0.05	-0.36**	-0.39**	-.30*	-.25*
3			–	0.01	-0.39**	-0.40**	-.48**	-.33**
4				–	-0.19*	-0.27*	-.20*	-.31**
5					–	0.38**	0.29*	0.20*
6						–	0.40**	0.24*
7							–	0.26*
8								–

German sample								
Discounting Inventory measure				Traditional discounting measure				
	1	2	3	4	5	6	7	8
1	–	0.40**	0.29*	0.19*	-0.51**	-0.39**	-0.36**	-0.29*
2		–	0.25*	0.16*	-0.46**	-0.35**	-0.30*	-0.23*
3			–	0.01	-0.42**	-0.39**	-0.50**	-0.26*
4				–	-0.30*	-0.32**	-0.28*	-0.39**
5					–	0.32**	0.48**	0.26*
6						–	0.39**	0.25*
7							–	0.27*
8								–

1. DI Delay discounting scale; 2. DI probability discounting scale; 3. DI effort discounting scale; 4. DI social discounting scale; 5. Traditional delay discounting; 6. Traditional probability discounting; 7. Traditional effort discounting; 8. Traditional social discounting.

* $p < .05$; ** $p < .01$

Davidshofer, 2005). This was assessed based on evidence of good internal consistency and factor analyses. It is also important to evaluate whether the Discounting Inventory measures the same construct as a traditional discounting task. Thus, the second goal of the study was to assess the construct validity of the DI by comparing discounting scores between this task and a widely used traditional discounting instrument.

Consistent with Malesza and Ostaszewski (in press), the four-factor solution gave the best model fit in both Polish and German samples. This solution was also supported by comparing four-factor model with one-factor model. These factors were closely associated with the theoretical four dimensions, which we have referred to as delay discounting, probabilistic discounting, social discounting, and effort discounting (Green & Myerson, 2004; Ostaszewski et al., 1998; Rachlin, 1993; Sugiwaka & Okouchi, 2004). The fit of the model is high when considering the χ^2/df or the RMSEA measures. According to the χ^2 statistic, however, the model would have to be rejected. This type of conflicting result is usually observed in personality models (Mulaik, 2007; Vassend & Skrandal,

1997). When using Chi square as a measure of model fit, the null hypothesis is that the model provides an adequate fit and thus a failure to reject the null indicates an adequate fit. The present analyses were based on samples of more than 200 participants each, so the Chi square tests were interpreted with caution. According to previous research, the sample size and the number of variables affect the χ^2 significance (Vassend & Skrandal, 1997). Thus, paying more attention to the χ^2/df measure is suggested. Furthermore, internal consistency coefficients were all acceptable (> 0.80) and to some point similar to those reported by Malesza and Ostaszewski (in press) in previous research. Thus, the results would suggest that the DI is a legitimate measure for use in the German population.

A secondary goal of the present study was to determine whether the strong relationship between scores on the DI and traditional discounting measures would be replicated in both samples of participants. In this case, construct validity was evaluated by comparing Discounting Inventory performance to that obtained using standardized delay, probability, social, and effort discounting tasks (Richards et al., 1999). Significant associations between

the two indicators of discounting were expected. On the basis of previously reported findings in the assessment of different measures of discounting (Malesza & Ostaszewski, in press; Reynolds, 2006; Reynolds & Schiffbauer, 2004; Richards et al., 1999), the author selected a correlation of .30 as the threshold criterion for establishing construct validity. Anastasi and Urbina (1997) suggest that such construct correlations should be “moderately high, but not too high” (p. 127), as very high correlations may suggest that the new measure is redundant. Groth-Marnat (2003) points out that there is no universally accepted minimal correlation sufficient to support construct validity. Rather, a criterion should be set logically, following the purpose and assumptions of the tests involved, and, where possible, comparison with known correlations among tests of the same construct.

Statistically significant correlations were detected between each DI scale and four traditional discounting measures. All of these correlations reached the criterion of 0.30. Each DI subscale correlated to 0.31 or better with its standard counterpart (except for the DI probability scale). However, delay and probability discounting exhibit considerable overlap. Thus, it can explain the stronger correlation between the DI probability construct with the traditional delay discounting measure than that between the DI probability scale and traditional probability instrument in the German sample. One reason for this effect might be that delay and probability reflect the same underlying decision-making process (for a review, see Green & Myerson, 2004), then one would predict a strong positive correlation between these two types of discounting. Indeed, prior research reported significant correlations between delay and probabilistic discounting (e.g., Richards, Zhang, Mitchell, & de Wit, 1999; Myerson, Green, Hanson, Holt, & Estle, 2003).

The consensus of the evidence suggests that the DI measures a similar construct to that measured by a traditional discounting instrument. However, based on the observed correlations, one might argue that the Discounting Inventory is a central conceptual component of impulsivity but only a peripheral component of discounting. Although a portion of this correlation could stem from common method variance, the datum still supports the assertion of construct correspondence between the two discounting measures. At one level of conceptualization, different instruments have varying degrees of overlap but no one measure is comprehensive (Reynolds et al., 2006). Even though one can argue that small to moderately significant correlations between these two discounting measures is due to the psychological construct *impulsivity*, the construct most commonly associated with discounting (Ainslie, 1975; Kirby & Finch, 2010; Reynolds et al., 2006).

Others (Mitchell, 1999; Richards et al., 1999) have interpreted the small association between self-reported scores and choice-based (behavioral) measures of discounting as evidence that the behavioral tendencies characterized by extreme discounting may not be the same as those indicated from self-report inventories of impulsivity – perhaps because assessments of discounting

isolate a more narrowly defined behavior (see Reynolds et al., 2006). Nonetheless, the results of the psychometric evaluation of the DI measure indicate that it meets standard, accepted criteria of construct validity and responsiveness for use as a discounting measure.

The research introduced in this report provides new information on validity of the DI. One potential limitation of the present study is the samples that were employed, that is university students. Discounting is associated with various maladaptive behaviors, and it is used as a diagnostic criterion for conditions such as various types of substance abuse (Madden, Petry, Badger, & Bickel, 1997) and pathological gambling (Petry, 2001). Thus, although one can conclude that the factor structure of the DI describes the present data quite well, one cannot conclude that the same factor structure would fit the data from a sample made up solely of clinical subjects. Given that the published research on the DI appears to support its construct validity, testing it in a clinical sample would appear to be the next logical step.

References

- Ainslie, G. (1975). Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin*, 82(4), 463–496.
- American Psychological Association (1992). Ethical principles of psychologists and code of conduct. *American Psychologist*, 47, 1597–1611.
- Anastasi, A., & Urbina, S. (1997). *Psychological testing (7th ed)*. Upper Saddle River, NJ: Prentice Hall.
- Arbuckle, J.L. (1997). *Amos users' guide version 3.6*. Chicago: Small Waters.
- Cronbach, L.J. (1990). *Essentials of psychological testing*. New York: Harper & Bros.
- Green, L., & Myerson, J. (1996). Exponential versus hyperbolic discounting of delayed outcomes: risk and waiting time. *American Zoologist*, 36, 496–505.
- Green, L., & Myerson, J. (2004). A discounting framework for choice with delayed and probabilistic rewards. *Psychological Bulletin*, 130, 769–792.
- Groth-Marnat, G. (2003). *Handbook of psychological assessment, 4th edition*. New York: John Wiley & Sons.
- Hambleton, R.K. (2001). The next generation of the ITC Test Translation and Adaptation Guidelines. *European Journal of Psychological Assessment*, 17(3), 164–172.
- Johnson, M.W., & Bickel, W.K. (2008). An algorithm for identifying non-systematic delay-discounting data. *Experimental Clinical Psychopharmacology*, 16, 264–274.
- Kirby, K.N., & Finch, J.C. (2010). The hierarchical structure of self-reported impulsivity. *Personality and Individual Differences*, 48, 704–713.
- Kirby, K.N., Petry, N.M., & Bickel, W.K. (1999). Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls. *Journal of Experimental Psychology: General*, 128(1), 78–87.
- Lei, M., & Lomax, R.G. (2005). The effect of varying degrees of non-normality in structural equation modeling. *Structural Equation Modeling*, 12(1), 1–27.
- Madden, G.J., & Bickel, W.K. (2010). *Impulsivity: The Behavioral and Neurological Science of Discounting*. APA Books, Washington, DC.
- Madden, G.J., Petry, N.M., Badger, G.J., & Bickel, W.K. (1997). Impulsive and self-control choices in opioid-dependent patients and non-drug-using control participants: drug and monetary rewards. *Experimental and Clinical Psychopharmacology*, 5, 256–262.
- Madden, G.J., Raiff, B.R., Lagorio, C.H., Begotka, A.M., Mueller, A.M., Hehli, D.J., & Wegener, A.A. (2004). Delay discounting of potentially real and hypothetical rewards: II. between- and within-subject comparisons. *Experimental and Clinical Psychopharmacology*, 12, 251–261.

- Malesza, M., & Ostaszewski, P. (in press). Discounting as a personality trait: Construction of the Discounting Inventory. *Frontiers in Psychology*.
- Mitchell, S.H. (2004). Effects of short-term nicotine deprivation on decision-making: Delay, uncertainty and effort discounting. *Journal of the Society for Research on Nicotine and Tobacco*, 6(5), 819–828.
- Mulaik, S. (2007). There is a place for approximate fit in structural equation modelling. *Personality and Individual Differences*, 42, 883–891.
- Murphy, K.R., & Davidshofer, C.O. (2005). *Psychological testing. Principles and applications: International edition*. Upper Saddle River, N.J.: Prentice-Hall, Inc.
- Myerson, J., & Green, L. (1995). Discounting of delayed rewards: Models of individual choice. *Journal of the Experimental Analysis of Behavior*, 64, 263–276.
- Myerson, J., Green, L., Hanson, J.S., Holt, D.D., & Estle, S.J. (2003). Discounting delayed and probabilistic rewards: Processes and traits. *Journal of Economic Psychology*, 24, 619–635.
- Myerson, J., Green, L., & Warusawitharana, M. (2001). Area under the curve as a measure of discounting. *Journal of the Experimental Analysis of Behavior*, 76, 235–243.
- Navarick, D.J. (2004). Discounting of delayed reinforcers: Measurement by questionnaires versus operant choice procedures. *The Psychological Record*, 54(1), 85–94.
- Ostaszewski, P. (1997). Temperament and the discounting of delayed and probabilistic rewards: Conjoining European and American Psychological traditions. *European Psychologist*, 2, 35–43.
- Ostaszewski, P., Green, L., & Myerson, J. (1998). Effects of inflation on the subjective value of delayed and probabilistic rewards. *Psychonomic Bulletin & Review*, 5, 324–333.
- Petry, N.M. (2001). Pathological gamblers, with and without substance use disorders, discount delayed rewards at high rates. *Journal of Abnormal Psychology*, 110, 482–487.
- Rachlin, H. (1993). The context of pigeon and human choice. *Behavior and Philosophy*, 21, 1–17.
- Rachlin, H., Raineri, A., & Cross, D. (1991). Subjective probability and delay. *Journal of the Experimental Analysis of Behavior*, 55, 233–244.
- Reynolds, B. (2006). The Experiential Discounting Task is sensitive to cigarette smoking status and correlates with a measure of delay discounting. *Behavioural Pharmacology*, 17, 133–142.
- Reynolds, B., Ortengren, A., Richards, J.B., & de Wit, H. (2006). Dimensions of impulsive behavior: Personality and behavioral measures. *Personality and Individual Differences*, 40, 305–315.
- Reynolds, B., & Schiffbauer, R. (2004). Measuring state changes in human delay discounting: an experiential discounting task. *Behavioural Processes*, 67, 343–356.
- Richards, J.B., Zhang, L., Mitchell, S., & de Wit, H. (1999). Delay and probability discounting in a model of impulsive behavior: effect of alcohol. *Journal of the Experimental Analysis of Behavior*, 71, 121–143.
- Smith, C.L., & Hantula, D.A. (2008). Methodological considerations in the study of delay discounting in intertemporal choice: A comparison of tasks and modes. *Behavior Research Methods*, 40(4), 940–953.
- Sugiwaka, H., & Okouchi, H. (2004). Reformative self-control and discounting of reward value by delay and effort. *Japanese Psychological Research*, 46, 1–9.
- Vassend, O., & Skrandal, A. (1997). Validation of the NEO Personality Inventory and the five-factor model. Can findings from exploratory and confirmatory factor analysis be reconciled? *European Journal of Personality*, 11, 147–166.

Appendix

Wir möchten Sie bitten eine Reihe von Fragen zu Ihren Verhaltensweisen und Persönlichkeitseigenschaften zu beantworten. Bei dieser Aufgabe gibt es weder richtige noch falsche Antworten, da jede Persönlichkeitseigenschaft seine Vorzüge hat. Die Ergebnisse dieses Fragebogens, werden ausschließlich zu Forschungszwecken verwendet, weshalb es besonders wichtig ist, dass die Antworten der Wahrheit entsprechen. Daher möchten wir Sie darum bitten sich zu jeder der unten stehenden Fragen zu positionieren, ohne darauf zu achten, was Sie in den vorherigen Fragen geantwortet haben. Bitte beschreiben Sie ehrlich wie Sie normalerweise sind und nicht wie Sie sich wünschen würden zu sein.

Sie müssen zu jeder der Folgenden Aussagen Bezug nehmen, indem Sie sich für jeweils eine der vier Antwortmöglichkeiten entscheiden.

- 1 – Ich stimme überhaupt nicht zu
- 2 – Ich stimme eher nicht zu
- 3 – Ich stimme eher zu
- 4 – Ich stimme voll und ganz zu

1. Normalerweise verzichte ich auf Dinge, welche ich nicht sofort haben kann.
2. Ich mag Glücksspiel
3. Allein die Tatsache, dass ich nicht willens bin mehr Anstrengung in meine Arbeit zu investieren, hält mich davon ab mehr zu erreichen.
4. Ich bevorzuge es, wenn andere wichtige Entscheidungen für mich treffen.
5. Im Eifer des Gefechts treffe ich Entscheidungen ungeachtet derer langfristigen Konsequenzen.
6. Ich würde lieber etwas ein wenig risikobehaftetes unternehmen (z.B. ein Auto über eine sehr anspruchsvolle kurvige Straße an einem steilen Abhang lenken) als ein paar Stunden ruhig und tatenlos zu verbringen.
7. Ich versuche es Aufgaben aus dem Weg zu gehen, welche körperliche Anstrengung erfordern.
8. Ich habe den Eindruck, dass andere Menschen mich immer übers Ohr hauen möchten.
9. Ich werde nervös, wenn ich lange anstehen muss.
10. Oft probiere ich neue Dinge nur aus, um den Nervenkitzel zu spüren.
11. Während körperlicher Anstrengung muss ich immer wieder Pausen einlegen um meine Kräfte zu regenerieren.
12. Ich kooperiere nur mit anderen, wenn ich weiß, dass es sich reichlich für mich auszahlt.
13. Ich höre oft, dass ich wie ein kleines Kind sei, weil ich Dinge immer sofort haben möchte.
14. Ich sehe keinen Grund dafür meine Arbeit fortzuführen, wenn keine große Chance besteht, dass sie sich lohnt.
15. Ich versuche so wenig wie möglich zu arbeiten, auch in Fällen wo andere mehr von mir erwarten.
16. Andere Menschen haben zu viel Kontrolle über mich.
17. Ich bevorzuge es Geld auszugeben statt es zu sparen.
18. Ich suche mit Absicht den Nervenkitzel, auch wenn es mich in Gefahr bringt.
19. Ich habe weniger Energie um etwas zu tun als andere.
20. Ich erwarte oft, dass andere meine Probleme für mich lösen.
21. Ich gebe oft spontan zu viel Geld aus und finde es schwierig zu sparen, sogar für besondere Ereignisse wie Urlaub.
22. Wenn ich an irgendeinem Spiel teilnehmen würde (z.B. Karten), würde ich es bevorzugen um Geld zu spielen statt nur zum Spaß.
23. Ich beende oft Aufgaben nicht, die ich anfangen.
24. Ich strebe normalerweise danach meine eigenen Bedürfnisse zu befriedigen, ohne andere zu beachten.
25. Mein Willen ist zu schwach um starken Versuchungen zu widerstehen.
26. Falls ich die Chance hätte schnell Geld zu verdienen würde ich in ein komplett neues Projekt investieren, wie eine neue und unbekannt Firma.
27. Ich bin schnell von Aufgaben gelangweilt, die nachdenken erfordern.
28. Ich verlasse mich immer auf die Meinung anderer.
29. Mich interessiert die Gegenwart mehr als die Zukunft.
30. Ich finde, es gibt keinen Spaß ohne Risiko, getreu dem Motto „no risk, no fun“.
31. Leicht werde ich entmutigt von Aufgaben, welche viel Arbeit mit sich bringen.
32. Ich versuche für gewöhnlich meine eigenen Bedürfnisse zu befriedigen, weil es unmöglich ist alle anderen zu befriedigen.
33. Ich versuche Aufgaben zu vermeiden, die keinen sofortigen Nutzen bringen.
34. Es kommt gelegentlich vor, dass ich per Anhalter fahre.
35. Pflichtaufgaben versuche ich mit so wenig Aufwand wie möglich zu bewältigen.
36. Ich ziehe es vor zu warten bis jemand anderes die Initiative ergreift um Aufgaben zu Ende zu bringen.
37. Ich bin nur daran interessiert was momentan passiert.
38. Emotionen lenken meine Entscheidungen.
39. Sicherlich könnte ich mehr erreichen, jedoch sehe ich keinen Grund dafür mehr zu tun als unbedingt nötig.
40. Einige denken ich sei geizig oder gierig
41. Ich will immer sofort die Ergebnisse meiner Arbeit sehen.
42. Ich glaube die Mehrheit der Menschen handelt überlegter und vorsichtiger als ich.
43. Meine schlechte Einstellung gegenüber meiner Arbeit, hält mich davon ab nutzbringende Resultate zu erzielen.
44. Ich kann nicht einmal meiner Familie oder Freunden vertrauen.
45. Ich kaufe oft Dinge, die ich in Wirklichkeit nicht brauche.
46. Ich würde eine gut bezahlte Stelle, die einfach zu verlieren ist, einer sicheren Stelle, die schlechter bezahlt ist, vorziehen.
47. Ich verliere schneller Energie während körperlicher Anstrengung als andere.
48. Menschen in meiner Umgebung zwingen mich oft dazu Entscheidungen gegen meinen Willen zu treffen.