

Other Papers

Polish Psychological Bulletin
2017, vol. 48(2) 293–300
DOI - 10.1515/ppb-2017-0033

Marta Malesza*
Maria Maczuga*

Reliability of the Discounting Inventory: An extension into substance-use population

Abstract: Recent research introduced the Discounting Inventory that allows the measurement of individual differences in the delay, probabilistic, effort, and social discounting rates. The goal of this investigation was to determine several aspects of the reliability of the Discounting Inventory using the responses of 385 participants (200 non-smokers and 185 current-smokers). Two types of reliability are of interest. Internal consistency and test-retest stability. A secondary aim was to extend such reliability measures beyond the non-clinical participant. The current study aimed to measure the reliability of the DI in a nicotine-dependent individuals and non-nicotine-dependent individuals. It is concluded that the internal consistency of the DI is excellent, and that the test-retest reliability results suggest that items intended to measure three types of discounting were likely testing trait, rather than state, factors, regardless of whether “non-smokers” were included in, or excluded from, the analyses (probabilistic discounting scale scores being the exception). With these cautions in mind, however, the psychometric properties of the DI appear to be very good.

Key words: Discounting Inventory, impulsivity, reliability, internal consistency, test-retest stability, personality trait

Introduction

In the personality literature, definitions of impulsivity vary widely and have included many concepts such as an inability to wait, an insensitivity to consequences, a tendency to risk taking, and cognitive and motor impulsivity (Ainslie, 1975; Kirby & Finch, 2010). It seems that the only consensus among researchers is that impulsivity is multidimensional in nature (Kirby & Finch, 2010; Reynolds, Ortengren, Richards, & de Wit, 2006). Given this controversy, 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

* Faculty of Psychology, University of Warsaw, Stawki 5/7, pok. 11, 00-183 Warszawa, Poland

Corresponding author: Marta Malesza

Acknowledgments

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).

value of the reinforcement diminishes according to the increasing number of people with whom the reward is to be shared; Rachlin, 1993).

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). 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). As has been seen, typical discounting measurement methods are somehow different from the usual forms of psychometric assessment. The presumption that questionnaire with (usually dozens) pairs of choices reflect behavioral processes requires critical examination (Navarick, 2004). Although widely used, some researchers have argued that the traditional discounting measure suffers from a number of practical problems. 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, recent 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). The primary reason for creating such a research tool was the need for a universal method of measuring individual differences in discounting that is independent of arbitrarily assumed types of rewards, delays, effort, etc.

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). The individual characteristics were selected separately for the delay, probabilistic, social, and effort discounting types and their aspects of behaviour. The generation of items and all consecutive steps were guided by methodological requirements underlying the construction of personality inventories (see Angleitner & Wiggins, 1986; Nunnally, 1978). The item content was related to the respective theoretical constructs of discounting (Rachlin, 1993). 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 from the Exploratory Factor Analyses, 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). Also, the correlational architecture of the DI corresponds well with what was observed using other traditional measures of the discounting rate. In particular, the correlations among the four types of discounting were mostly weak, as postulated by previous studies of discounting (see Madden & Bickel, 2010). 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). Those 48 items had loadings equal to or higher than 0.40 on their own factor, and lower on the remaining factors. 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: $r = 0.20-0.47$). The final 48-item version of the inventory seems to be a relevant alternative to the traditional discounting measures.

While such an instrument shows great promise as a research and clinical assessment tool, almost no data has yet been collected regarding the fundamental psychometric properties of the instrument. The most fundamental of these properties is reliability, which establishes the upper bound for validity (Anastasi & Urbina, 1997). Two types of reliability are of interest. Internal consistency refers to the extent to which items comprising the scale measure the same construct (i.e., homogeneity of scale), and is assessed by Cronbach's α (Cronbach, 1990) at any given administration of the test. However, recent work suggests that test-retest reliability may be more predictive than estimates of internal consistency (Murphy & Davidshofer, 2005). Test-retest reliability, an indicator of the stability of a measuring instrument, is assessed by administering the instrument to respondents on two different occasions and examining the correlation between test and retest scores. The length of the test-retest interval should be short enough to ensure that clinical change in the symptom being measured is unlikely to occur, but sufficiently long to ensure that respondents do not recall their responses from first assessment (Cronbach, 1990).

Furthermore, instruments are particularly useful if they can measure significant change in behavior indicators over time (e.g., pre–vs. post–intervention/treatment). Results from pre–post research designs are only credible to the extent that the measure used is not prone to error across administrations. That is, only to the extent that apparent effects are not due to score instability (Murphy & Davidshofer, 2005). To our knowledge, one reliability study has been conducted on the DI method (Malesza & Ostaszewski, in press).

Previous research on reliability of discounting

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 $p < 0.05$. The internal consistencies of the DI measured with Cronbach’s alphas are also adequate. The following coefficients for the internal consistency of the scales were observed: total measure $\alpha = 0.89$, effort discounting scale $\alpha = 0.95$, probabilistic discounting scale $\alpha = 0.88$, social discounting scale $\alpha = 0.82$, and delay discounting scale $\alpha = 0.87$, suggesting that the items have relatively high internal consistency. Hence, these findings suggest that all of these subscales are reasonable, indicating good internal consistency and test–retest stability of the DI measure. Malesza and Ostaszewski (in press) concluded that the internal consistency of the DI was good, and that the test–retest reliability results suggested that items intended to measure four types of discounting were likely testing trait, rather than state, factors. Thus, there is evidence that discounting is reasonably stable over modest time frames and with different assessment techniques (Odum, 2011). Scientifically speaking, the construction of the DI leads to a hypothesis that the discounting rate can be regarded as an individual personality trait. Stability over time is one of the defining characteristics of personality traits (Murphy & Davidshofer, 2005).

Compared to the DI, the test–retest reliability of the traditional discounting measure has been well documented. Test–retest reliability remains good up to intervals of one year (Baker, Johnson, & Bickel, 2003; Beck & Triplett, 2009; Kirby, 2009; Ohmura, Takahashi, Kitamura, & Wehr, 2006; Simpson & Vuchinich, 2000; Smith & Hantula, 2008; Takahashi, Funikawa, Miyakawa, Maesato, & Higuchi, 2007). However, most of these studies employed non-clinical subjects; only one study reported moderate to high test–retest correlations for the traditional discounting scores in a substance-abuse population comparing smokers and nonsmokers (Baker et al., 2003).

Substance Use and Discounting

Previous research showed that steep discounting appears to be a common phenomenon in a range of populations with different maladaptive behavioral issues (for a review, see Madden & Bickel, 2010). A growing literature highlights important associations between discounting patterns and substance use and abuse.

Substance-dependent individuals have been found to discount hypothetical monetary gains at a higher rate than matched controls. Indeed, a large literature clearly shows that individuals who are dependent on or abuse nicotine (e.g., Reynolds, Richards, Horn, & Karraker, 2004), alcohol (e.g., Petry, 2001), heroin (e.g., Madden, Petry, Badger, & Bickel, 1997), and cocaine (e.g., Heil, Johnson, Higgins, & Bickel, 2006), evidence steeper rates of discounting (i.e., are more impulsive) than do non-drug-using controls. One of the major criticisms of research on the DI to date is that it has been conducted using participants who do not qualify as potential substance-use individuals (Malesza & Ostaszewski, in press). This fact is problematic given that the DI was designed also to be used with this population. The current study comprehensively investigated the reliability of the discounting outcomes in both current cigarette smokers and never-before cigarette smokers.

Overview

While traditional discounting measures are known to produce stable measures of discounting across re-testing intervals ranging up to one year (e.g., Kirby, 2009; Ohmura et al., 2006; Simpson & Vuchinich, 2000) little is known about the test–retest reliability of the DI. Malesza and Ostaszewski (in press) reported that DI discounting rates did not differ significantly when reassessed 2 weeks later. However, the short re-test interval raises the possibility that choices made in the first session influenced those made in the second session. Thus, the goal of the present investigation was to test the reliability of the DI after an interim of 6 months. The hypotheses of this investigation was that the internal consistency and test–retest reliability would equal or surpass those observed for the previous study on the DI (Malesza & Ostaszewski, in press), and that the test–retest reliability of the DI would compare favorably to that of the traditional discounting measures. The 6-month interval was chosen for the following reason. It was assumed that the 6-month retest time would provide a more stringent test of the DI’s temporal stability. Test–retest reliability tends to decrease over time (Cronbach, 1990). Thus, if the DI’s temporal stability were “good” at 6 months, it would be reasonable to assume that it would be “good” or “very good” at one or three months.

No study, however, has tested the reliability of the DI measure in substance-dependent individuals (Malesza & Ostaszewski, in press). To fill this gap, we administered the DI to a sample of current-cigarette smokers and examined its psychometric properties and factor structure. We predicted that DI would retain its sound psychometric properties, thus paving the way for its use in current-smokers sample and facilitating research on other substance-dependent individuals across the discounting process.

Participants

Individuals were recruited from the university and surrounding community through advertisements posted in public area on campus and on social media (i.e., Facebook).

The total sample consisted of 385 participants (157 female and 228 male; 200 non-smokers and 185 smokers) ranging in age from 18 to 47 years ($M = 22.85$, $SD = 2.46$) but with approximately 67% of them aged between 19 and 29 years. Most of them were high school graduates or had a higher level of education (65% high school, 28% university degree). Fifty-eight percent of participants were students at the time of data collection. A total of 385 individuals completed the DI a second time, 6 months following its initial administration. During this period current-smokers did not take part in any therapeutic (cessation) activities leading to quit smoking which could influence the way they filled out the questionnaire. All participants provided written informed consent after the nature of the study had been explained to them. 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).

Potential participants were asked a series of questions to determine if they qualified. Current smokers were people who reported smoking at least 10 cigarettes per day, had a Fagerström Test for Nicotine Dependence (Heatherton, Kozlowski, Frecker, & Fagerstroem, 1991) score of at least 6, and answered “yes” to at least three questions on the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV; American Psychiatric Association, 1994) drug checklist for cigarettes. Never-before smokers were people who reported that they had never smoked cigarettes. People who met these qualifications were invited for additional testing. Other study exclusion criteria included the following: a history of psychiatric diagnosis; current use of any stop-smoking treatments; regular use of illegal drugs or other forms of tobacco (the exception being nicotine dependence for current cigarette smokers); or an inability or unwillingness to complete study activities.

Measures

At baseline, participants completed questionnaires regarding demographics, smoking history, and cigarette dependence. The Fagerström Test of Cigarette Dependence (FTCD; Heatherton et al., 1991) is a widely-used six-item self-report measure of physical dependence on cigarettes. Higher scores (range 0–10) signify higher rates of physical dependence on nicotine. The internal consistency of the FTCD is fair (Cronbach’s $\alpha = 0.61$; Heatherton et al., 1991). Internal consistency was slightly higher in the current study (Cronbach’s $\alpha = 0.68$). Next, participants received the Discounting Inventory (Malesza & Ostaszewski, in press). The measurement comprised the 48-item pool with the 4-point Likert scale format (4 = fully agree, 3 = agree slightly, 2 = disagree slightly, 1 = disagree completely). Six months after the initial administration, all individuals from the original sample completed the ID a second time. In both the original and retest conditions, participants were tested individually and responses remained anonymous.

Results

Internal Consistency Reliability

Internal consistency reliability statistics (Cronbach’s α) were calculated for never-before smokers, current-smokers, and all respondents from both subsamples, for the overall DI scores (48 items), and for each of the content scores: delay, probability, social, and effort discounting scales (twelve items each; no items shared between the four content scores). Results are displayed in Table 1.

For the total sample ($N = 385$, both current-smokers and never-before smokers included), internal consistency was excellent. Reliability for the total DI score was 0.92. Reliability was high for the content components, ranging from 0.83 (delay discounting scale) to 0.88

Table 1. Alpha coefficients and correlations across test sessions 1 and 2

Scale	Cronbach’s α Session 1			Cronbach’s α Session 2			Pearson-r Session 1 & 2*		
	Current-smokers	Non-smokers	All participants	Current-smokers	Non-smokers	All participants	Current-smokers	Non-smokers	All participants
Delay Discounting	0.79	0.90	0.83	0.91	0.87	0.93	0.87	0.88	0.82
Probabilistic Discounting	0.86	0.80	0.81	0.83	0.82	0.90	0.59	0.67	0.57
Effort Discounting	0.89	0.91	0.88	0.92	0.85	0.95	0.79	0.84	0.80
Social Discounting	0.85	0.84	0.87	0.88	0.86	0.94	0.85	0.82	0.78
Overall discounting	0.87	0.88	0.92	0.89	0.90	0.93	0.79	0.85	0.85

* all correlations significant at $p < 0.001$; α = Cronbach’s alpha

(effort discounting scale). When data from never-smokers individuals were excluded (current smokers $N = 185$) and the coefficients were recalculated, internal consistency for the overall DI ($\alpha = 0.87$), and content score reliability (ranged from 0.79 for delay discounting scale to 0.89 for effort discounting scale) remained high. For subsample 2 ($N = 200$ of non-smokers; current-smokers excluded) overall internal consistency was high ($\alpha = 0.88$). Among the content scores, effort discounting scale yielded the best reliability ($\alpha = 0.91$), and probabilistic discounting scale produced the lowest ($\alpha = 0.80$).

Internal consistency was also calculated for the 385 participants who completed the DI a second time 6 months after the original administration. Alpha scores for this second administration were high for the DI overall ($\alpha = 0.93$), as well as for the items measuring each scale (α from 0.90 to 0.95). Again, when the data from the non-smokers participants were excluded and the coefficients recalculated, internal consistency measures remained high for the DI overall ($\alpha = 0.89$), as well as for the four content scores ranged from 0.83 to 0.92. When the data from the current-smokers subsample were excluded, overall internal consistency was 0.90, and the coefficients for the content scores ranged from 0.82 (probabilistic discounting scale) to 0.87 (delay discounting scale).

Test–retest Reliability

Of the original sample, all participants ($N = 385$) completed the DI again after a delay of 6 months. Test–retest reliability was determined by calculating Pearson correlation coefficients between the parameters obtained during Session 1 and the parameters obtained during Session 2 for the overall score on the 48-item DI, as well as for each scale separately. Table 1 presents these correlations for the current and never-before smokers, and for all participants. Results were considered significant at $p < 0.05$ and all reported correlations were significant at $p < 0.001$, two-tailed.

Total DI score temporal stability was adequate for the whole sample (0.85). Test–retest coefficients for

the component scores were generally fair to adequate, ranging from 0.78 (social discounting scale) to 0.82 (delay discounting scale). The notable exception was the probabilistic discounting score (0.57). A second analysis was conducted on data from participants who currently smoke. For these participants, test–retest reliability was 0.79 for the overall DI score, and content scores ranged from 0.59 for probabilistic discounting scores to 0.87 for delay discounting scores. The comparison test–retest reliability for the never-smokers was 0.85, whereas scores on each scale correlated at 0.67 (probabilistic discounting) to 0.88 (delay discounting; see Figure 1).

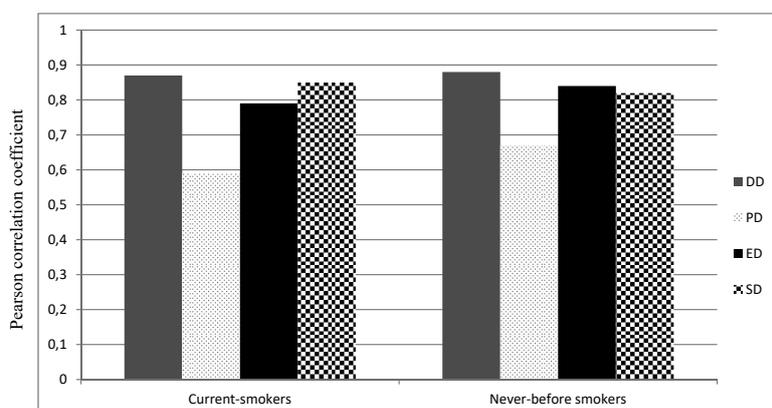
Current-smokers versus never-before smokers

Current-smokers smoked on average 104 cigarettes (SD = 12 cigarettes) a week. These comparisons imply that Fagerström Test of Cigarette Dependence ratings were strongly correlated with the rate of smoking (i.e., number of cigarettes smoked per week). The correlation between the two measures was $r = 0.84$, $p < 0.001$. The additional aim for using two subsamples was to increase the chances of detecting differences between current-smokers and never-before smokers. Comparison of the overall discounting score between current-smokers and never-before smokers was conducted with Mann–Whitney U test. Analysis determined that the discounting parameter for current-smokers differed significantly from the discounting parameter for never-before smokers ($U = 112$, $p = 0.01$). As predicted current smokers discounted at higher rates than did never-before smokers.

Confirmatory Factor Analyses

To examine the factor structure of the data, we performed the Confirmatory Factor Analyses with the AMOS statistical package (Arbuckle, 1997). We used several criteria of model fit (see Mulaik, 2007). A well-fitting model should ideally have a nonsignificant χ^2 statistic ($p > 0.05$), the GFI (goodness-of-fit index), TLI (Tucker–Lewis index), and CFI (comparative fit index) values close to 0.95 or greater. The model is considered to

Figure 1. Pearson correlation coefficients between parameters obtained during Session 1 and the parameters obtained during Session 2



Note. DD = delay discounting; PD = probabilistic discounting; ED = effort discounting; SD = social discounting.

have a reasonable fit if the GFI, TLI, and CFI values are approximately 0.90. As the χ^2 statistic tends to be inflated in small samples, 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 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. We note, however, that the choice of indices and cutoff values is a topic surrounded by considerable controversy (see, e.g., Mulaik, 2007).

The analysis for the never-before smokers subsample showed that the four factor model had a good fit to the data according to the goodness of fit indicators: $\chi^2(80) = 147.25$, $p > 0.05$, $\chi^2/df = 1.84$, RSMEA = 0.05, GFI = 0.95, CFI = 0.98, TLI = 0.99. Confirmatory Factor Analyses for the current-smokers participants showed that the four factor model had an acceptable fit to the data $\chi^2(125) = 250.00$, $p = 0.04$, $\chi^2/df = 2.00$, RSMEA = 0.05, GFI = 0.90, CFI = 0.96, TLI = 0.97. In both subsamples, the four factor model had a much better fit than a one factor model of general discounting ($\Delta\chi^2_{(55)} = 770.1$, $p < 0.001$ for current-smokers; $\Delta\chi^2_{(29)} = 513.14$, $p < 0.001$ for never-before smokers).

Discussion

Reliability is a necessary but not sufficient aspect of validity and must be assessed before determining that a test is valid (Anastasi & Urbina, 1997). The goal of this investigation was to determine several aspects of the reliability of the Discounting Inventory (Malesza & Ostaszewski, in press). A secondary aim was to extend such reliability measures beyond the non-substance-use participant. While the DI was developed based on the research on non-substance-dependent population, applied research often requires comparison with substance-dependent individuals. Thus, psychometric data gleaned from this population will be necessary in order to establish the test reliability with respect to its intended use. The current study aimed to measure the internal consistency and test–retest reliability of the DI in a substance-use (nicotine-dependent individuals) and non-substance-use (non–nicotine-dependent individuals) samples.

Internal Consistency

In terms of internal consistency, the DI performed extremely well regardless of whether “non-smokers” were included in, or excluded from, the analyses. The lowest observed Cronbach’s alpha was 0.79. As a point of comparison, these results were superior to those reported for the original DI (see Malesza & Ostaszewski, in press), where the consistency of respondents’ discounting rates across items ranged in proportion from 0.82 to 0.95.

Test–retest Reliability

The DI test–retest reliability indicators were more mixed. The total DI score ($rtt = 0.85$) fell well within the limits of acceptable reliability (Groth-Marnat, 2003). Reliability varied by content scales. Test–retest coefficients for content scores were generally poorer than for the total DI scores. However, stability was the same for current-smokers (from 0.59 to 0.87) and for never-before smokers (0.67 to 0.88). These data indicate that a significant positive correlation was observed for scores obtained during two sessions measuring discounting for each type. For both subsamples (substance-use and non-substance-use), the lowest correlation was observed for probabilistic discounting scale and the highest correlation was observed for delay discounting scale. Furthermore, to put these results in perspective, it should be noted that the DI retained respectable test–retest coefficients relative to comparable traditional discounting measures, despite the much longer test–retest intervals experienced by the DI respondents. As a part of comparison, moderate reliability was demonstrated after six weeks (Beck & Triplett, 2009), whereas weak correlations were noted across quarters (Kirby et al., 2002). Also, previous DI retest scores were highly reliable at a 2-week interval (Malesza & Ostaszewski, in press). The present study used time intervals of 6 months between sessions, potentially increasing the ecological validity of the findings.

The pattern of reliability coefficients was particularly interesting for the probabilistic scale scores. Discounting shows clear trait influences (Odum, 2011). The previous research verified the status of the discounting rate as a trait (Malesza & Ostaszewski, in press) and brings arguments supporting the hypothesis that discounting can be considered as a stable personality trait. However, probability scores evidenced a much lower test–retest reliability than did the other components, indicating that the probability scale scores changed significantly as a function of time, whereas other component scores (and the overall score) were comparatively stable. These data suggest that probability scale scores may reflect more of a state than trait function in maintaining discounting process.

Confirmatory Factor Analyses

Confirmatory factor analyses have been performed to test the adequacy of the structure of this model. The analysis replicated the four-factor structure postulated by Malesza and Ostaszewski (in press), but that the model was only acceptable for the current-smokers subsample. Note that the RMSEA, CFI, TLI, and χ^2/df all indicated a good model fit but the GFI measure was low. 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 & Skrondal, 1997). According to previous research, the small sample size, if we consider the number of variables in each model, may also contribute to the weakness of these results. Thus, paying more attention to the χ^2/df measure is suggested (Vassend & Skrondal, 1997). According to this measure, the four-factor model has a reasonable fit.

Current-smokers versus never-before smokers

We also tested the prediction that never-before-smokers will show smaller discounting than current-smokers in the discounting procedure. Higher rates of discounting seem to be prevalent in populations that are typically described as impulsive, such as substance-dependent individuals (see Mackillop et al., 2011 for review and meta-analysis). Therefore, discounting may serve as a useful index of impulsivity. The results of the present study revealed differences in discount scores between current cigarette smokers and never-before cigarette smokers. Current-smokers' discount scores were higher than never-before smokers' rates. These findings are comparable to and consistent with the literature, which have demonstrated that substance-using individuals have a greater tendency to discount rewards than controls with less or no substances use (e.g., Baker et al., 2003).

Second, reliability of the current-smokers population was not significantly lower than those observed for the never-before smokers. Both the current and never-before smokers' discount scores were stable. These findings suggest that individual differences in impulsive behaviors measured by the DI are stable and reproducible enough. This finding is informative and important with respect to extending the utility of the DI as a clinical instrument, supporting the utility and feasibility of using the DI for the assessment of impulsive behaviors in substance-use population. Thus, nicotine-dependent individuals, just like non-nicotine-dependent individuals, may have stable discounting scores. Although nicotine-dependent individuals may, in general, discount outcomes at a higher rate than non-nicotine-dependent individuals (e.g., Baker et al., 2003), their discount scores did not change in time.

However, additional research will be necessary to determine if the DI performs equally well on the broader population for which it has been designed. It is the case that the present investigation utilized a homogenous population in terms of age (i.e., mainly young people), and occupation (i.e., mainly university students). Conclusions based on the present results should take these factors into account. It is also worthy of mention that the present study assessed only one type of substance abuse (cigarettes). Thus, one cannot assume that the present reliability results are representative for all types of addictions. Making such a determination would seem a worthy pursuit for future research.

With these cautions in mind, however, the psychometric properties of the DI appear to be very good. In sum, the DI's internal consistency coefficients and test-retest reliability not only meet common professional standards (Groth-Marnat, 2003), but are also comparable to those of popular measures of discounting rate (e.g., Kirby, 2009; Ohmura et al., 2006; Simpson & Vuchinich, 2000). As noted previously, there are a number of different techniques for measuring discounting, each with strengths and weaknesses. The strengths of the DI method are that participants provide exact identification of potential mechanism that underlie the discounting rate, and data collection is rapidly completed. In such instances, this method may be the most realistic. Both researchers and therapists using typical screening measures might find it

beneficial to also employ the DI because the resulting data might highlight why that respondents are characterized by steep discounting. Now that discounting can be studied using a variety of methods, researchers should consider the various strengths and weaknesses of each when deciding which method to employ.

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.
- American Psychiatric Association. (1994). Diagnostic and statistical manual of mental disorders: DSM-IV [Internet]. 4th ed. Washington (DC): American Psychiatric Association.
- Anastasi, A., & Urbina, S. (1997). *Psychological testing (7th ed)*. Upper Saddle River, NJ: Prentice Hall.
- Angleitner, A., & Wiggins, J.S. (1986). *Personality assessment via questionnaires. Current issues in theory and measurement*. Berlin–Tokyo: Springer-Verlag.
- Arbuckle, J.L. (1997). *Amos users' guide version 3.6*. Chicago: Small Waters.
- Baker, F., Johnson, M.W., & Bickel, W.K. (2003). Delay discounting in current and never-before cigarette smokers: Similarities and differences across commodity, sign, and magnitude. *Journal of Abnormal Psychology*, 112, 382–392.
- Beck, R.C., & Triplett, M.F. (2009). Test-retest reliability of a group-administered paper-pencil measure of delay discounting. *Experimental and Clinical Psychopharmacology*, 17, 345–355.
- 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.
- Heatherington, T., Kozlowski, L., Frecker, R., & Fagerström, K. (1991). The Fagerström test for nicotine dependence: A revision of the Fagerström tolerance questionnaire. *British Journal of Addiction*, 86, 1119–127.
- Heil, S.H., Johnson, M.W., Higgins, S.T., & Bickel, W.K. (2006). Delay discounting in currently using and currently abstinent cocaine-dependent outpatients and non-drug-using matched controls. *Addictive Behaviors*, 31, 1290–1294.
- Kirby, K.N. (2009). One-year temporal stability of delay-discount rates. *Psychonomic Bulletin and Review*, 16, 457–462.
- Kirby, K.N., & Finch, J.C. (2010). The hierarchical structure of self-reported impulsivity. *Personality and Individual Differences*, 48, 704–713.
- Kirby, K.N., Godoy, R., & et al. (2002). Correlates of delay-discount rates: Evidence from Tsimane' Amerindians of the Bolivian rain forest. *Journal of Economic Psychology*, 23, 291–316.
- MacKillop, J., Amlung, M.T., Few, L.R., Ray, L.A., Sweet, L.H., & Munafò, M.R. (2011). Delayed reward discounting and addictive behavior: a meta-analysis. *Psychopharmacology*, 216(3), 305–321.
- 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.
- Navarick, D.J. (2004). Discounting of delayed reinforcers: Measurement by questionnaires versus operant choice procedures. *The Psychological Record*, 54(1), 85–94.
- Nunnally, J.C. (1978). *Psychometric theory*. New York: McGraw-Hill.
- Odum, A.L. (2011). Delay discounting: Trait variable? *Behavioural Processes*, 87, 1–9.
- Ohmura, Y., Takahashi, T., Kitamura, N., & Wehr, P. (2006). Three-month stability of delay and probability discounting measures. *Experimental and Clinical Psychopharmacology*, 14, 318–328.
- 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). Delay discounting of money and alcohol in actively using alcoholics, currently abstinent alcoholics, and controls. *Psychopharmacology*, 154, 243–250.
- 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., 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., Richards, J., Horn, K., & Karraker, K. (2004). Delay and probability discounting as related to cigarette smoking status in adults. *Behavioural Processes*, 65, 35–42.
- Simpson, C.A., & Vuchinich, R.E. (2000). Reliability of a measure of temporal discounting. *The Psychological Record*, 50, 3–16.
- 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.
- Takahashi, T., Furukawa, A., Miyakawa, T., Maesato, H., & Higuchi, S. (2007). Two-month stability of hyperbolic discount rates for delayed monetary gains in abstinent inpatient alcoholics. *Neuroendocrinology Letters*, 28, 131–136.
- 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.