Published in final edited form as:
Exp Clin Psychopharmacol. 2014 October ; 22(5): 434–443. doi:10.1037/a0037340.

Predictive Validity of Delay Discounting Behavior in Adolescence: A Longitudinal Twin Study

Joshua D. Isen, Ph.D., **Jordan C. Sparks, M.A.**, and **William G. Iacono, Ph.D.**
Department of Psychology, University of Minnesota, Twin Cities

Abstract

A standard assumption in the delay discounting literature is that individuals who exhibit steeper discounting of hypothetical rewards also experience greater difficulty deferring gratification to real-world rewards. There is ample cross-sectional evidence that delay discounting paradigms reflect a variety of maladaptive psychosocial outcomes, including substance use pathology. We sought to determine whether a computerized assessment of hypothetical delay discounting (HDD) taps into behavioral impulsivity in a community sample of adolescent twins (N = 675). Using a longitudinal design, we hypothesized that greater HDD at age 14–15 predicts real-world impulsive choices and risk for substance use disorders in late adolescence. We also examined the genetic and environmental structure of HDD performance. Individual differences in HDD behavior showed moderate heritability, and were prospectively associated with real-world temporal discounting at age 17–18. Contrary to expectations, HDD was not consistently related to substance use or trait impulsivity. Although a significant association between HDD behavior and past substance use emerged in males, this effect was mediated by cognitive ability. In both sexes, HDD failed to predict a comprehensive index of substance use problems and behavioral disinhibition in late adolescence. In sum, we present some of the first evidence that HDD performance is heritable and predictive of real-world temporal discounting of rewards. Nevertheless, HDD might not serve as a valid marker of substance use disorder risk in younger adolescents, particularly females.

Keywords

delay discounting; impulsivity; substance use; heritability

It is self-evident that, all else being equal, a currently available reward is more valuable and appealing than a future reward. However, situations often arise in which the deferral of immediate gratification facilitates the eventual attainment of larger rewards. Individuals who adaptively exploit these investment opportunities are rightfully considered more patient and intelligent (Burks, Carpenter, Götte, & Rustichini, 2009). Under laboratory conditions, experimenters can administer scenarios that directly gauge individuals' competing preferences between consumption and investment. The ability to forgo immediate consumption in order to receive larger future rewards is referred to as delay of gratification.

Correspondence concerning this article should be addressed to Joshua Isen, Department of Psychology, University of Minnesota, 75 East River Road, Minneapolis, MN 55455. jdisen@umn.edu, Phone: (612) 625-1535, Fax: (612) 626-2079.

Disclosures
The authors do not have any conflicts of interest in relation to the present research and its implications.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

A complementary phenomenon is delay discounting (DD), which describes the willingness to forfeit future gains in order to secure more immediate advantages. As the expected delivery of a reward increases in temporal remoteness, so does the willingness to accept a lesser, but imminently available, alternative.

Although theoretically complementary, the concepts of delay of gratification and DD are emphasized quite differently in the literature (Steinberg et al., 2009).[1] The tasks that measure delay of gratification typically involve a single choice between an immediate reward and a larger delayed one. Crucially, participants experience real-time reinforcement, i.e., immediate reward choices are reinforced immediately, and delayed rewards are delivered after the relevant delay has elapsed. Under this paradigm, failure to endure the required delay to receive the larger, more desirable reward is interpreted as reflecting poorer resistance to temptation, and thus inferior impulse control (Mischel, Shoda, & Rodriguez, 1989). Conventional DD tasks, by contrast, involve multiple trials to determine one's discount rate. The value of the immediate reward is systematically manipulated in order to arrive at an indifference amount – the point at which the immediate reward and the future reward are subjectively equivalent. Delay intervals are typically wide-ranging (e.g., varying from one day to one year), such that preference for the immediate reward option is not inherently maladaptive. For practical purposes (including time and budgetary constraints), reward choices are not continuously reinforced, and thereby must be regarded as essentially hypothetical.

Evidence concerning the association between hypothetical DD performance and self-reported impulsivity is equivocal. This is consistent with the multi-faceted nature of impulsivity and the general observation that laboratory tasks which purportedly measure impulsivity often show little overlap with self-report measures (Reynolds, Penfold, & Patak, 2008; Kirby & Finch, 2010). It is not clear which particular aspects of impulsivity, if any, are related to DD behavior (Reynolds, Ortengren, de Wit, & Richards, 2006; but see de Wit, Flory, Acheson, McCloskey, & Manuck, 2007). According to a recent large study, DD was found to be quite distinct from personality measures that tap into self-control (Steinberg et al., 2009). Instead, DD was linked to a particular facet that the authors refer to as future orientation (e.g., "Should I get a job after high school or enroll in college and go into debt?"). The latter may be independent from the broader construct of impulsivity.

Nonetheless, DD appears to be a robust correlate of behaviors that ostensibly stem from poor self-control. It has been related to a variety of social maladies in adolescents, including academic failure, substance use, and antisocial behavior (e.g., Reynolds & Fields, 2012; Chesson et al., 2006; Field, Christiansen, Cole, and Goudie, 2007). A recent meta-analysis demonstrates that discounting rates are steeper in populations manifesting addictive behavior (MacKillop et al., 2011). Effect sizes between individuals with and without

[1]Technically, the term "delay of gratification" (DG) refers to a specific paradigm within the research tradition of willpower (e.g., Stanford Marshmallow Experiment), in which delayed reward choices must be sustained during an interval of some (unforeseeable) length. According to this definition, a necessary challenge in DG tasks is to restrain one's impulses in the face of ongoing temptation, or to otherwise defect (Reynolds & Schiffbauer, 2005). One can entertain a broader definition of DG (e.g., refusing a tantalizing and immediately consumable reward), in which the opportunity to defect (change one's mind) is not an available option. While the *decision* to defer gratification would arguably qualify as (absence of) delay-discounting behavior, we wish to emphasize that standard delay-discounting tasks do not typically involve a tantalizing, immediately consumable reward.

substance dependence are typically of moderate magnitude, and show minimal heterogeneity when restricted to multi-trial DD tasks. Moreover, steeper DD is associated with the broader externalizing spectrum (Bobova, Finn, Rickert, & Lucas, 2009); DD performance is related to the common variance shared by different externalizing disorders rather than to any specific substance-related characteristic.

Another strand in the literature has emphasized the inverse association between cognitive ability and steepness of DD. A meta-analysis of two dozen studies, including five hypothetical-reward and eight experiential DD measures, reveals a correlation of $r = -.25$ between discount rate and IQ test performance (Shamosh & Gray, 2008). Delay-related behavior does not appear to preferentially relate to verbal ability over general intelligence. As a result, low IQ may potentially confound any observed associations between disinhibitory psychopathology and DD (Wilson, Mitchell, Musser, Schmitt, & Nigg, 2011). Rather than representing a distinct component of impulsivity, DD performance might profitably be viewed as an aspect of general intelligence.

Present Study

The overarching goal of this study is to examine the predictive validity of DD in typically developing adolescents. To our knowledge, only two previous studies have examined prospective associations between delay discounting behavior and future substance use (Audrain-McGovern et al., 2009; Fernie et al., 2013); the vast majority of investigations have been cross-sectional. Adolescence is a period in which individuals typically begin experimenting with addictive substances, and is thus an opportune time to determine whether DD behaviors predict substance use pathology. Using a longitudinal twin cohort, we administered one of two delay-related tasks at mid-adolescence (14–15 years old) and late adolescence (17–18 years old). At the younger assessment wave, we administered a computerized DD task involving hypothetical monetary scenarios. At the follow-up assessment (approximately three years later), we administered a one-shot “Cash Choice” paradigm that utilizes real-world cash incentives (Sparks, Isen, & Iacono, 2014). The delayed reward amount ($10) was constant across the two tasks.

We used several approaches to determine whether hypothetical delay discounting (HDD) predicts behavioral impulsivity. First, we examined whether HDD discount rates are prospectively related to the Cash Choice outcome. If delay-related impulsivity is a developmentally stable disposition, then HDD and Cash Choice should show significant congruence with one another. Secondly, we evaluated its predictive validity with respect to externalizing psychopathology and self-reported impulsivity. We examined intelligence test scores and socioeconomic status as potential confounds for any emergent HDD-Externalizing associations. By assessing both verbal and nonverbal intellectual abilities, we sought to determine whether HDD represents a component of the broader construct of intelligence.

A secondary objective is to ascertain the genetic and environmental contributions to HDD performance. Some authors have speculated that the genes influencing DD behavior also underlie genetic risk for substance use disorders, i.e., DD is an endophenotype for substance use problems (Audrain-McGovern, Nigg, & Perkins, 2009; Mitchell, 2011). At a very minimum, this would require that DD behavior exhibit significant heritability. Several studies have isolated candidate genes that influence DD (e.g., Paloyelis, Asherson, Mehta, Faraone, & Kuntsi, 2010; Smith & Boettiger, 2012), but none have used a classic twin design to examine the degree to which individual differences in HDD rates are heritable. The relevant behavioral genetics literature has focused solely on the one-shot Cash Choice paradigm (Anokhin, Golosheykin, Grant, & Heath, 2011). Sparks et al. (2014) demonstrated that additive genetic influences account for approximately one-half of the Cash Choice variability in the present sample. We suspect that standard HDD tasks are more sensitive to inter-individual variability, as the underlying measures (indifference points) are continuously distributed rather than dichotomous.

Method

Participants

Participants were drawn from a population-based sample of twins from the state of Minnesota. The twins were first recruited at age 11 to take part (with their caregivers) in a comprehensive day-long laboratory visit. They participated in the Enrichment Study component of the Minnesota Twin Family Study (Keyes et al., 2009), which investigates the antecedents of substance use disorders. The twins returned for follow-up assessments at approximate three-year intervals. As of this writing, the second follow-up assessment is complete. The sample was “enriched” for externalizing problems during recruitment, insomuch that half of the twins were selected for disruptive behavior problems and academic disengagement. They were screened for these signs via a phone interview with a parent. The remaining participants were randomly selected from an unscreened sample of twins born between 1988 and 1994. These groups were combined to form the Enrichment Study. The recruitment strategy and sampling characteristics are described in considerable detail by Keyes et al. (2009).

One thousand twins participated at intake when they were 11–12 years old (M = 11.9, SD = 0.43). All participants were members of same-sex twin pairs, and included 478 males and 522 females. The sample was 91% Caucasian. Zygosity determination was based on an algorithm of sibling resemblance across various anthropometric indices (cephalic index, fingerprint ridge count, and ponderal index) as well as subjective evaluations of physical similarity. Molecular genetic confirmation was used to resolve any discrepancies between these estimates. The ratio of monozygotic (MZ) to dizygotic (DZ) twins was 3:2. Most participants (> 85%) returned for two follow-up visits in adolescence. Their mean ages at the first follow-up (FU1) and second follow-up (FU2) visits were 15.1 (SD = 0.55) and 17.9 (SD = 0.46), respectively.

Measures

Delay discounting of hypothetical monetary rewards (HDD)—A computerized HDD task was administered at FU1. It was added to the protocol after the assessment wave had already begun, such that 675 of 930 FU1 participants received this task. Task length and trial sequence differed across participants due to our employment of a random adjusting-amount procedure (Richards, Zhang, Mitchell, & de Wit, 1999; see below). Participants were faced with two options at each trial. They had the choice of receiving an immediate reward or a standard $10 payment after some future delay. Immediate rewards spanned from $0.50 to $10.00 and were adjusted in increments of $0.50. There were six delay intervals: 1 day, 2 days, 10 days, 30 days, 180 days, and 365 days.

The combination of delays and immediate reward amounts was presented in a pseudo-random order using the E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA). An immediate reward amount was presented randomly to participants on the opening trial of each delay interval. The program adjusted the amount of the immediate reward on successive trials in order to arrive at an “indifference point” in which the immediate amount is subjectively equivalent to the delayed $10 reward. Participants’ responses to previous trials were used to narrow the range of values from which the amount of each subsequent trial was selected. Upper and lower bounds were established to ensure that responses systematically converged upon an indifference point (see Richards et al., 1999). If a response was inconsistent with previous responses, then the program would not converge upon an indifference point, and the relevant range of values (i.e., upper limit or lower limit) was reset. Thus, the total number of trials was dependent on participants’ performance. Participants who responded in a diligent and consistent manner required fewer trials to establish an indifference amount at each delay interval. Given that reward values were not presented in a predictable manner (e.g., in ascending or descending order), it was common for participants to show inconsistent preferences. However, these errant responses posed little harm, as the program would probe participants for additional information if an inconsistency emerged.

Participants did not receive actual payment contingent on their responses. Instead, they were instructed to imagine receiving the particular reward they selected on each trial. For example, if they chose to receive $5 now instead of $10 in 30 days, then they were told to pretend as if they would receive $5 in cash at the end of the session. If their preference was for the delayed reward, then they were instructed to imagine laboratory staff putting the $10 in an envelope with their name and address on it. Then it would be sent to them after 30 days.

After instructions were displayed on the computer screen, the participant was given two practice trials. Laboratory staff supervised the participant during these practice trials to make certain that he/she understood the nature of the task, and queried the participant if necessary. The recorded task session began after the conclusion of the practice trials.

Cash Choice paradigm—At the conclusion of the second follow-up visit (FU2), we administered a one-shot delay discounting paradigm devised by Wulfert, Block, Santa Ana, Rodriguez, and Colsman (2002) to 791 twins.[2] Participants were given the choice between receiving $7 in cash instantly or a $10 bill to be mailed within a week. [Through pretesting,

[2]Wulfert et al. (2002) characterize this task as a measure of self-regulation that appropriately extends the “delay-of-gratification” paradigm to adolescents. Given that participants are not required to sustain their willpower (i.e., there is no possibility of later succumbing to temptation and reversing one’s decision), this paradigm is more precisely a measure of “experiential” delay-discounting. Nonetheless, in some respects, this paradigm is more similar to DG tasks than to conventional (multi-trial) DD tasks.

Wulfert et al. (2002) determined that these particular reward options ($7 vs. $10) were sufficiently close in value to generate conflicting preferences in youth.] Co-twins were placed in separate rooms, and were unaware of each other's choice. If the delayed choice was selected, then participants were instructed to write their mailing address on an envelope. Laboratory staff directly interacted with participants during this task, and orally presented the following instructions:

> We are interested in how people make decisions, so we are going to ask you to make a choice. As a token of our appreciation for your participation today, we'd like to offer you a cash bonus. You can either choose to have $7 right now [showing cash to the participant] or wait. If you wait, we'll mail $10 [showing the cash] to you. You should receive it in approximately one week. What would you like to do?

Trait impulsivity—We obtained self-report ratings of impulsivity from an abbreviated version of the Multidimensional Personality Questionnaire (MPQ; Tellegen & Waller, 2008) at both FU1 and FU2 assessments. The MPQ scale of interest is Control (vs. Impulsivity). Previous work indicates that it is highly correlated with total scores from the Barratt Impulsiveness Scale (Critchfield, Levy, & Clarkin, 2004), especially with motor and non-planning aspects of impulsiveness (Vaidya, Latzman, Markon, & Watson, 2010). Our MPQ Control scale contains 18 items that tap into the broad impulsivity construct (e.g., "When faced with a decision, I usually take time to consider and weigh all options".) Each item was rated on a four-point scale – *definitely false* (1), *probably false* (2), *probably true* (3), and *definitely true* (4). Mean scores were calculated only for participants who responded to at least 16 items; 98.7% of participants (N = 906) had valid data at FU1, and all but two individuals had valid data at FU2 (N = 884). The alpha coefficients were 0.86 and 0.87 at FU1 and FU2, respectively. MPQ Control showed good rank-order stability across the approximate three-year interval, Pearson's r = .58. Since mean levels of self-control slightly increased over time (d = 0.12), we averaged the z-scores for these two assessments. Thus, our impulsivity index is based on combining the scores from FU1 and FU2. (Results from our inferential statistical analyses were virtually identical when we examined the scores from each assessment separately.)

Intelligence—A short form of the Wechsler Adult Intelligence Scale-Revised (Wechsler, 1981) was administered to participants at FU2. We used the Vocabulary and Block Design subtests. These particular subtests were selected because they assess both verbal and nonverbal domains, and are highly correlated with full-scale IQ (Sattler, 2001). Raw scores were converted to scaled score equivalents, which are age-standardized.

Socioeconomic Status—Socioeconomic status (SES) was a latent variable composed of household income, highest parental occupational status, maternal educational attainment, and paternal educational attainment. Data were obtained from the intake assessment during the years 1999–2005. Household income was construed as an ordinal variable corresponding to ten income brackets: less than $20,000, more than $100,000, and eight intermediate deciles. Parental occupational status was based on seven categories from a Hollingshead index. We selected the highest parental occupational status (rather than using both parental indicators) because employment decisions of one spouse might be influenced by the co-spouse's occupational status. Educational attainment refers to each parent's highest academic degree. All four indicators cohered to a great extent, with loadings on the SES factor spanning from .62 to .77. Factor scores were estimated in Mplus version 6 (Muthen & Muthen, 2010).

Externalizing Psychopathology—We measured individuals' liability for externalizing psychopathology at FU2 using lifetime criteria of substance use and antisocial behavior. We also measured substance use at FU1 in order to capture early-onset behaviors that are more contemporaneous with participants' HDD performance. This allows us to determine whether delay discounting predicts a future progression of substance use involvement (see Audrain-McGovern et al., 2009; Fernie et al., 2013). Since symptoms of substance dependence are extremely uncommon at age 14, we conceived our FU1 substance use measure as dichotomous. Namely, we recorded whether individuals had yet tried nicotine, alcohol, or cannabis at their FU1 assessment. Given that access to alcohol is often sanctioned by the family (e.g., sipping of wine during religious rites), we focused on whether participants had ever consumed alcohol *without* parental permission. Based on confidential responses to a computerized substance use survey, a total of 27.6% participants had used nicotine, cannabis, or alcohol (without parental consent) prior to FU1.

At FU2, we obtained a comprehensive measure of externalizing liability (EXT) that indexes alcohol consumption frequency, alcohol dependence, illicit substance use/dependence, nicotine use/dependence, and behavioral disinhibition. All substance-related information was assessed through the Substance Abuse Module of the Composite International Diagnostic Interview (Robins, Baber, & Cottler, 1987). Our latent EXT score was based on a higher-order factor derived from 17 facet measures and five intermediate factors. A detailed description of the various indicators (as well as a rationale for their selection) can be found in Hicks et al. (2011). The structure of this model and its accompanying factor loadings are presented in Sparks et al. (2014). The higher-order EXT factor accounted for at least 50% of the variance in each of the five intermediate factors. (A confirmatory factor analysis yielded good fit, as the root mean square error of approximation was .045.) After extracting factor scores for each participant in Mplus, we used a normalizing (Blom) transformation to obtain a normal distribution of EXT.

Discounting analysis

We quantified HDD using an area-under-the-curve (AUC) approach (Myerson, Green, & Warusawitharana, 2001), which sums the trapezoidal areas bounded by pairs of indifference points at adjacent delay intervals. This method was chosen because it does not rely on a priori assumptions about the distribution of indifference points. In particular, it makes no assumption that the data points conform to a nonlinear shape (e.g., hyperbolic or exponential).

Higher values of AUC represent the cumulative effects of possessing larger indifference amounts. Thus, individuals with lower AUC values are more impatient, i.e., exhibit greater HDD behavior. Appropriate quantification of HDD is hindered, however, if participants show erratic preferences across the different delay intervals. In theory, indifference amounts should consistently decline as the delay until reward delivery increases in length. Nonsystematic patterns of indifference amounts would indicate that participants are responding in an unreflective manner. We adapted an algorithm devised by Johnson and Bickel (2008) to identify individuals with questionable data. Nonsystematic data were flagged whenever the indifference amount at a given delay exceeded that of the immediately preceding (shorter) delay by a magnitude of $2.50 or more. We also flagged cases if we observed any cumulative increase ≥ $2.50 over the course of multiple successive delays (wherein no single adjacent increase exceeded $2.00). All delay discounting data from such participants were considered invalid.[3]

Data are also suspect if participants fail to exhibit discounting behavior due to their insensitivity to delay. Johnson and Bickel (2008) recommend exclusion of cases in which the initial and final indifference amounts (e.g., 1-day vs. 365-day delay) differ by less than 10%. We deemed such a criterion too strict because the maximum delay interval was merely a year in our study, whereas the aforementioned investigators reviewed tasks featuring delays as long as 25 years. Instead, we were only concerned with participants who selected the delayed reward on *every* trial. These participants are potentially problematic because their responses are too systematic and perhaps not reflective of their actual preferences. Ultimately, their data were retained for all analyses because we could empirically demonstrate that these delay-insensitive individuals are more patient and intelligent than other participants. Our substantive results did not change when we excluded them.

Statistical Analyses

For all statistical analyses, we accounted for the non-independence of observations by computing standard errors using a sandwich estimator in Mplus. Otherwise, standard errors around the means are biased in twin samples where observations are repeated measures within the pair.

We examined the genetic and environmental structure of delay discounting by fitting quantitative genetic models to the AUC data. Maximum likelihood estimation was used to compare three empirically-driven structural equation models. Our baseline model specified that twin covariance was due to a combination of additive genetic and non-additive (e.g., dominance) effects. This model is appropriate if the MZ twin correlation is more than twice the size of the DZ magnitude. (Non-additive genetic effects occur when alleles interact in a dominant/recessive manner or if the effects of a given set of genes depend on the actions of other genes. With non-additive genetic influences, trait resemblance will be markedly reduced in DZ pairs. Additive genetic effects occur when the products of alleles independently "add up" to determine the trait.) The expected twin covariance matrices were based on standard assumptions in biometric model-fitting (Neale & Cardon, 1994), with an additional parameter representing nonshared environmental effects. In sum, additive and non-additive genetic correlations were fixed to 0.50 and 0.25, respectively, in DZ twins,

[3]In response to a reviewer's suggestion, we ran inferential statistical analyses without invalidating these participants' data. Compared to when their data were excluded, the magnitudes of the association between HDD and criterion variables (e.g., substance use) were weakened. Effect sizes were always pulled in a direction that was contrary to theoretical expectation.

whereas they were set to unity in MZ twins. (DZ twins share half of their alleles on average – and a quarter of all allelic interactions at any given locus – whereas MZ twins share 100% of their inherited genetic material.) In the next two models, we separately constrained each of these genetic parameters to zero. The model that yielded the lowest Bayesian Information Criterion (BIC) was selected as best-fitting.

Results

Descriptive Statistics

Means and standard errors for all measures are listed in Table 1. In order to ascertain whether there were significant sex differences in the mean level of each variable, we conducted a series of linear regression models and a probit model (in the case of Cash Choice threshold). Members of both sexes strongly preferred the delayed $10 reward over the immediate $7 reward during the Cash Choice paradigm. Although females chose the delayed reward in a slightly greater proportion than males (80% vs. 73%), this difference did not reach statistical significance, $p = .06$. Intelligence test scores and SES were similar across sex. The prevalence of early-onset substance use (by age 14) was nearly identical across the sexes. However, by age 17–18, males exhibited greater lifetime EXT liability, $p < .01$, and also reported lower self-control scores (combined across ages 14–15 and 17–18), $p = .04$.

Validation of HDD data

Of the 675 participants who received the HDD task, 3.9% ($n = 26$) lacked sufficient data because their responses were repeatedly inconsistent (thereby preventing the computer program from successfully converging upon indifference points.) (The E-prime software was programmed to terminate after a given number of convergence attempts failed.) Data from these individuals were excluded from all analyses. Of the remaining participants, Cronbach's alpha indicated relatively high internal consistency among the six indifference point values ($\alpha = 0.86$). However, this conceals the fact that 92 participants (14.2%) qualitatively showed nonsystematic data across the different delay lengths. We identified a final group of 48 participants who failed to show DD behavior because they always selected the $10.00 reward.

We compared the two problematic groups of participants (i.e., those with nonsystematic data and those who were systematically delay-insensitive) to the larger group of participants with valid data ($n = 509$). Dummy codes were used to represent these groups, with normal discounters treated as the reference group. First, we conducted a probit regression model predicting the Cash Choice outcome. Relative to the normal discounting group, systematic delayers had an excess 17.4% likelihood of selecting the delayed cash reward, but this difference was non-significant, $p = .08$. Participants with nonsystematic data had a 15.1% reduction in likelihood of selecting the delayed reward, $p = .02$. Next, we ran a series of linear regression models predicting external criteria. Systematic delayers had higher block design scores relative to the normal discounting group, $p < .01$, $d = 0.51$. Otherwise, no significant group differences emerged for vocabulary score, SES, self-reported impulsivity, and EXT (all *p*s > 0.16), Most importantly, individuals with nonsystematic data had higher raw AUC values than the normal discounting group, $p < .01$, $d = 0.66$. Inclusion of the former thus introduces considerable bias, as it misleadingly indicates that they are markedly more patient than the latter. This bias does not apply to the systematic delayers, however. Systematic delayers showed a greater trend of selecting the delayed Cash Choice reward, suggesting that they behave much like individuals who are extremely patient, i.e., rightfully placed at the extreme end of the DD distribution.

Discount rates were similar between males and females. Figure 1 depicts the average indifference amounts in each sex (after excluding participants with nonsystematic data). These data points visually conform to a hyperbolic function. The AUC values were 0.38 and 0.40 in males and females, respectively. For subsequent analyses, AUC was converted from its raw form to rank-normalized (Blom) values to yield a more normal distribution. Blom transformed AUC values ranged from −2.67 to 1.81, with a standard deviation of 0.98. Age was not significantly correlated with AUC, Pearson's $r = .06$.

Quantitative Genetic Analyses

Given that there was no sex difference in AUC, we were able to combine males and females when examining the twin covariance. Means and variances of AUC could be equated across twin birth order and zygosity without significant loss of fit; $p = .55$, BIC = 1550. The correlation coefficient in 148 pairs of MZ twins was $r = 0.44$ (*SE* = .07), whereas the correlation in 97 pairs of DZ twins was small and non-significant, $r = .14$ (*SE* = .10). This observed discrepancy in twin resemblance is consistent with a model where individual differences in HDD are due to genetic and nonshared environmental influences. Structural equation modeling revealed that the optimal genetic parameter was non-additive in nature; our best-fitting model did not retain the additive genetic component ($p = .72$, BIC = 1539). Maximum-likelihood estimation returned a heritability estimate of .47 (*SE* = .06), i.e., 47% of the inter-individual variability in HDD can be attributed to non-additive genetic influences.

Predictive Validity

Next we examined whether HDD performance is associated with real-world temporal discounting of rewards. We were interested in whether AUC as well as the underlying indifference amounts would prospectively relate to participants' Cash Choice. Given that the Cash Choice task featured a week-long delay, we hypothesized that the indifference amount at the 10-day delay would show the closest correspondence. In order to examine this possibility, we computed polychoric correlations between the Cash Choice outcome and ordered polytomous variables corresponding to each indifference amount. (Since the latter were not symmetrically distributed, it was necessary to recode each indifference amount into ten ordinal categories, typically by rounding values to the nearest whole dollar.)

The correlations among Cash Choice, AUC, and the six indifference amounts are reported in Table 2. There was a modest correlation between Cash Choice and AUC. Cash Choice outcome was significantly associated with the 10-day indifference amount as well as those at larger delay intervals. All six indifference amounts were significantly correlated with one another, polychoric *r*s > .53 (data not shown).

The construct validity of HDD was further examined using relevant external criteria, including intelligence test scores, SES, early-onset substance use (Early SU), lifetime EXT liability, and MPQ Control. Correlations among these variables (see Table 3) are reported separately in males and females in order to detect the possibility of sex-specific effects, although the overall pattern was quite similar between the sexes. Hypothetical DD performance (AUC) was largely unrelated to the three report-based measures of behavioral disinhibition (i.e., MPQ Control, Early SU, and EXT). A single exception occurred in males, in which the correlation between AUC and Early SU use was significant ($p = .01$). Male substance users exhibited greater HDD relative to non-users by a 0.37-unit difference in AUC, Cohen's $d = -0.38$. Given that this association trended in the opposite direction in females, the interaction of sex (females = 0, males = 1) with substance use group was significant ($b = -0.47$, $p = .02$). As expected on the basis of previous work (Sparks et al., 2014), Cash Choice showed robust associations with disinhibitory psychopathology and self-reported impulsivity.

Another pattern in Table 3 is that HDD consistently related to both verbal ability (Vocabulary) and nonverbal reasoning (Block Design). As a result, we examined whether the link between Early SU and HDD performance in males persists after controlling for cognitive ability. The association was reduced to non-significance when including either Vocabulary or Block Design as a covariate in the model. Finally, we tested whether cognitive ability mediates the association between HDD performance and Cash Choice. In a multivariate probit regression model, we examined whether AUC, Vocabulary, and Block Design uniquely contributed to the likelihood of selecting the delayed cash-choice reward. To increase statistical power, we combined the male and female samples. All three variables significantly predicted Cash Choice, *p*s < .02. Thus, HDD behavior accounts for unique variance in real-world temporal discounting (i.e., independent of intelligence).

Discussion

There has been considerable enthusiasm in the clinical utility of DD, with recent proclamations that DD might serve as an endophenotype for substance use disorders (Anokhin et al., 2011; Mitchell, 2011; Audrain-McGovern et al., 2009). As such, DD performance should reflect a heritable disposition that is tied to a broader network of impulsivity. We were able to confirm that individual differences in HDD, as measured by area-under-the-curve, are moderately heritable in adolescent twins. However, HDD is not a valid model of substance use disorder risk in the present sample.

The preponderance of literature indicates that individuals with substance dependence exhibit greater DD (i.e., lower AUC values) as compared to healthy individuals (MacKillop et al., 2011). We therefore hypothesized that HDD performance would be related to externalizing psychopathology in our adolescent sample. This prediction was weakly supported in males, insomuch that steeper delay discounting was related to a past history of substance use. Hypothetical DD in both sexes failed to prospectively relate to a latent EXT score indexing substance use/dependence and behavioral disinhibition. In this respect, our results contradict those of Fernie et al. (2013), in which delay discounting performance in early adolescence predicted future progression of alcohol use problems.

Despite its lack of relationship with trait impulsivity and disinhibitory psychopathology, HDD performance predicted impulsive outcomes three years later on the Cash Choice task. This is a novel finding, given that previous studies have mostly failed to detect significant overlap between HDD and real-word (experiential) delay discounting rates (e.g., Melanko & Larkin, 2013; Paloyelis et al., 2010). This incongruence appears to result from the vastly greater delay lengths and reward magnitudes typically employed by hypothetical paradigms as compared to experiential ones (Scheres, Sumiya, & Thoeny, 2010). We expected to find substantial correspondence between participants' choices on the two tasks, given that the magnitude of the delayed reward ($10) was identical. Moreover, the Cash Choice paradigm features a delay (7 days) that is similar to one of the delay intervals (10 days) used in the HDD task. As a one-shot assessment, our Cash Choice paradigm is not comparable to conventional (multi-trial) experiential DD tasks, in which delays are necessarily limited in length – measured in seconds rather than days (Hinvest & Anderson, 2010; Melanko & Larkin, 2013; Reynolds & Schiffbauer, 2004; Scheres et al., 2010).

Both the HDD and Cash Choice paradigms were related to verbal and nonverbal intelligence scores. This is in accordance with the available literature on DD and cognitive ability (Shamosh & Gray, 2008). It is interesting that performance on the DD tasks was related to Block Design no less than Vocabulary. This serves to distinguish delay-related impulsivity from externalizing behaviors and self-reported impulsivity, in which the principal correlate is verbal ability (Isen, 2010). Remarkably, females' HDD behavior was just as highly correlated with Block Design as the two intelligence measures were associated amongst themselves. Taken together, these phenomena suggest that HDD performance taps into general intellectual ability.

Heritability of HDD was in the 40–50% range, which is similar to estimates obtained from the Cash Choice task (Sparks et al., 2014). Although HDD and Cash Choice reveal similar broad-sense heritability estimates, the underlying genetic mechanisms may differ. We found that non-additive genetic factors were likely responsible for twin resemblance in HDD, whereas genetic influences on Cash Choice appear to be additive in nature. This is illustrated by the fact that the DZ twin concordance for Cash Choice outcome is moderate in magnitude ($r = 0.35$; Sparks et al., 2014), approaching that of MZ twins, and suggestive of family-wide environmental influences. In contrast, the DZ correlation for HDD is small ($r = 0.14$), but this should be interpreted with caution, as it was based on fewer than 100 twin pairs. More twin studies and molecular genetic investigations are needed to corroborate whether genetic dominance and epistatic mechanisms influence HDD behavior.

Assessing the Construct of Impulse Control

In the present sample, the Cash Choice paradigm appears to more optimally tap into the reward immediacy and behavioral undercontrol that standard HDD tasks purportedly reflect. It is often assumed that the hypothetical preferences manifested on standard DD tasks mirror real-world behavioral tendencies, but this assumption is not necessarily tenable (Hinvest &

Anderson, 2010; Melanko & Larkin, 2013). Task conditions in which rewards are hypothetical (or "potentially real") do not engage the same appetitive processes that are triggered by the salience and immediacy of real-world rewards (Navarick, 2004). For example, Scheres, Lee, and Sumiya (2008) found that self-ratings of hyperactivity/ impulsivity in college students were related to DD rates when real money was at stake but not when rewards were hypothetical.

The hot/cool-system model of Metcalfe and Mischel (1999) is germane to this discussion, as it offers a theoretical framework for understanding the nature of impulse control. The "hot" system deals with automatic responses to appetitive stimuli, whereas the "cool" system translates such stimuli into mental abstractions. In the context of delay of gratification, individuals who implement cool representations are better able to distance themselves from the "hot" stimulus, thereby facilitating the execution of long-term goals (Metcalfe & Mischel, 1999). The Cash Choice task likely activates the hot system because it uses a reward that is both highly visible and motivating. Participants must resist the tempting wad of dollar bills – at least for a moment. On the other hand, the HDD task employs content that is strictly informative-cognitive, leading to less visceral engagement. In the absence of temptation, there is little need for the cool system to inhibit hot system processing, i.e., impulse control is not involved. This difference may account for the distinct validities of our HDD and Cash Choice tasks.

Current conceptualizations of the externalizing spectrum propose that diverse externalizing characteristics – including substance dependence and antisocial behavior – cohere due to a common vulnerability to poor impulse control (Iacono, Malone, & McGue, 2003; Hicks, Krueger, Iacono, McGue, & Patrick, 2004). However, in order for impulse control to be directly engaged (e.g., when craving a cigarette or contemplating shoplifting), there must first be activation of the hot system. If HDD performance was indeed unaffected by hot system processing in the present sample, then this might account for its lack of relationship with self-reported impulsivity and externalizing psychopathology.

Limitations and Future Directions

A major limitation is that the HDD and Cash Choice tasks were not administered contemporaneously. This renders it difficult to assess their relative utility. The correspondence between these two tasks was almost certainly attenuated due to the three-year gap between assessments. Moreover, there are developmental changes in self-control competencies between mid-adolescence and late adolescence. Notwithstanding these issues, we believe it is unlikely that the two tasks assess the same construct.

A further limitation is that Cash Choice showed limited variability in our sample; nearly 80% of participants selected the delayed reward. This result is unexpected, especially given the fact that a slightly older sample of college freshmen failed to show an overall preference when presented with these same choices (Wulfert et al., 2002). Future studies may profit from adjusting the parameters (e.g., extending the delay length or increasing the immediate reward amount) in order to maximize variability in the choices of young adults. On the other hand, there is always the attendant risk that the association between "cash choice" and other measures (such as HDD) will change unpredictably after adjusting these task parameters.

Another objection is that standard HDD tasks may be less suitable for use in children due to the greater abstraction of the scenarios. It is possible that task validity was undermined due to the immaturity of our participants. Previous research has demonstrated that young children do not show systematic delay discounting (Reynolds & Schiffbauer, 2005). However, preadolescence appears to be the turning point at which individuals begin manifesting sensitivity to different delays. As our participants were 14–15 years of age, they should have been sufficiently competent to perform in an adaptive, consistent manner. Indeed, the curves in Figure 1 reveal that participants, as a whole, showed adequate discounting behavior.

A potential weakness is that we used a single task to operationalize HDD – one that utilizes an adjusting-amount procedure. Our results may not generalize to studies that have employed other formats, as discounting curves are not necessarily interchangeable across different tasks (Epstein et al., 2003). Nonetheless, the paradigm described by Richards et al. (1999) is quite standard in the literature. Numerous studies have reported significant differences in DD between individuals with and without substance dependence using this procedure (e.g., Fields, Leraas, Collins, & Reynolds, 2009; Reynolds & Fields, 2012; also see MacKillop et al., 2011). It is unclear why such findings are so ubiquitous in the literature, while we were unable to reproduce a similar effect using a much larger sample size. An admittedly speculative explanation is that we used a nonclinical sample whereas other studies explicitly sampled individuals with substance use problems. Our failure to demonstrate an association between HDD performance and externalizing problems is consistent with the results of Olson, Cooper, Collins, and Luciana (2007), in which community samples of children and young adults were used.

A nontrivial number of participants possessed nonsystematic data, which might undercut task validity. This phenomenon appears to be a consequence of the random adjusting-amount procedure, in which combinations of delays and reward amounts are presented in an unpredictable order. Consequently, participants cannot rely on their short-term memory to ensure that responses are consistent across trials. We were compelled to exclude data from 17.5% of participants due to qualitative patterns of nonsystematic responding. Such a high degree of attrition is not unique to our sample. Studies that have used this procedure in youth have reported comparable rates of nonsystematic responding (e.g., 17% in Wilson et al., 2011; 18% in Olson et al., 2007). Our group of nonsystematic responders showed greater impatience on the Cash Choice task, but there was no evidence that they suffered greater disinhibitory psychopathology.

Our results raise questions about possible sex and age-specific contributions to the association between delay discounting and substance use. Sex moderated the association between HDD behavior and early-onset substance use. This is a novel finding that requires further corroboration. There are some hints of sex and age moderation in a large-scale study of youth reported by Romer, Duckworth, Sznitman, and Park (2010). The relationship between HDD performance and substance use was stronger in males and in participants aged 18–22 (as compared to females and those aged 14–17), although the interactions were not significant. Further research is necessary to explore whether demographic variables such as sex and age affect HDD-substance use relationships.

Acknowledgments

This research was supported by NIH grant DA 013240. The role of this funding source was solely financial. Joshua Isen's effort on this project was supported by T32 grant MH017069 from the National Institute of Mental Health.

All three authors significantly contributed to the composition of this manuscript. Joshua Isen was assisted by Jordan Sparks in conducting statistical analyses and drafting the manuscript. William Iacono provided detailed feedback and helped edit the text.

We would like to acknowledge Dr. Stephen M. Malone for his instrumental role in software programming and data processing. We also wish to acknowledge Jennifer M. Donnelly for overseeing task administration.

References

Anokhin AP, Golosheykin S, Grant JD, Heath AC. Heritability of delay discounting in adolescence: a longitudinal twin study. Behavior Genetics. 2011; 41:175–183. [PubMed: 20700643]

Audrain-McGovern, J.; Nigg, JT.; Perkins, K. Endophenotypes for nicotine dependence risk at or before initial nicotine exposure. In: Swan, GE.; Baker, TB.; Chassin, L.; Conti, DV.; Lerman, C.; Perkins, KA., editors. Phenotypes and Endophenotypes: Foundations for Genetic Studies of Nicotine Use and Dependence (Chapter 8). National Cancer Institute Monograph, No. 20, US Department of Health & Human Services, National Institutes of Health; 2009. p. 339-402.

Audrain-McGovern J, Rodriguez D, Epstein LH, Cuevas J, Rodgers K, Wileyto EP. Does delay discounting play an etiological role in smoking or is it a consequence of smoking? Drug and Alcohol Dependence. 2009; 103:99–106. [PubMed: 19443136]

Bobova L, Finn PR, Rickert ME, Lucas J. Disinhibitory psychopathology and delay discounting in alcohol dependence: personality and cognitive correlates. Experimental and Clinical Psychopharmacology. 2009; 17:51–61. [PubMed: 19186934]

Burks SV, Carpenter JP, Götte L, Rustichini A. Cognitive skills affect economic preferences, strategic behavior, and job attachment. Proceedings of the National Academy of Sciences of the United States of America. 2009; 106:7745–7750. [PubMed: 19416865]

Chesson H, Leichliter J, Zimet G, Rosenthal S, Bernstein D, Fife K. Discount rates and risky sexual behaviors among teenagers and young adults. Journal of Risk and Uncertainty. 2006; 32:217–230.

Critchfield KL, Levy KN, Clarkin JF. The relationship between impulsivity, aggression, and impulsive-aggression in borderline personality disorder: an empirical analysis of self-report measures. Journal of Personality Disorders. 2004; 18:555–570. [PubMed: 15615667]

de Wit H, Flory JD, Acheson A, McCloskey M, Manuck SB. IQ and nonplanning impulsivity are independently associated with delay discounting in middle-aged adults. Personality and Individual Differences. 2007; 42:111–121.

Epstein LH, Richards JB, Saad FG, Paluch RA, Roemmich JN, Lerman C. Comparison between two measures of delay discounting in smokers. Experimental and Clinical Psychopharmacology. 2003; 11:131–138. [PubMed: 12755457]

Fernie G, Peeters M, Gullo MJ, Christiansen P, Cole JC, Sumnall H, Field M. Multiple behavioural impulsivity tasks predict prospective alcohol involvement in adolescents. Addiction. 2013; 108:1916–1923. [PubMed: 23795646]

Field M, Christiansen P, Cole J, Goudie A. Delay discounting and the alcohol Stroop in heavy drinking adolescents. Addiction. 2007; 102:579–586. [PubMed: 17309540]

Fields S, Leraas K, Collins C, Reynolds B. Delay discounting as a mediator of the relationship between perceived stress and cigarette smoking status in adolescents. Behavioral Pharmacology. 2009; 20:455–460.

Gibson HB. Self-reported delinquency among schoolboys, and their attitudes to the police. The British Journal of Social and Clinical Psychology. 1967; 6:168–173. [PubMed: 6077466]

Hicks BM, Krueger RF, Iacono WG, McGue M, Patrick CJ. Family transmission and heritability of externalizing disorders: a twin-family study. Archives of General Psychiatry. 2004; 61:922–928. [PubMed: 15351771]

Hicks BM, Schalet BD, Malone SM, Iacono WG, McGue M. Psychometric and genetic architecture of substance use disorder and behavioral disinhibition measures for gene association studies. Behavior Genetics. 2011; 41:459–475. [PubMed: 21153693]

Hinvest NS, Anderson IM. The effects of real versus hypothetical reward on delay and probability discounting. The Quarterly Journal of Experimental Psychology. 2010; 63:1072–1084. [PubMed: 19890767]

Iacono WG, Malone SM, McGue M. Substance use disorders, externalizing psychopathology, and P300 event-related potential amplitude. International Journal of Psychophysiology. 2003; 48:147–178. [PubMed: 12763572]

Isen J. A meta-analytic assessment of Wechsler's P>V sign in antisocial populations. Clinical Psychology Review. 2010; 30:423–435. [PubMed: 20350774]

Johnson MW, Bickel WK. An algorithm for identifying nonsystematic delay-discounting data. Experimental and Clinical Psychopharmacology. 2008; 16:264–274. [PubMed: 18540786]

Keyes MA, Malone SM, Elkins IJ, Legrand LN, McGue M, Iacono WG. The enrichment study of the Minnesota twin family study: increasing the yield of twin families at high risk for externalizing psychopathology. Twin Research and Human Genetics. 2009; 12:489–501. [PubMed: 19803776]

Kirby KN, Finch JC. The hierarchical structure of self-reported impulsivity. Personality and Individual Differences. 2010; 48:704–713. [PubMed: 20224803]

MacKillop J, Amlung MT, Few LR, Ray LA, Sweet LH, Munafo MR. Delayed reward discounting and addictive behavior: a meta-analysis. Psychopharmacology. 2011; 216:305–321. [PubMed: 21373791]

Melanko S, Larkin KT. Preference for immediate reinforcement over delayed reinforcement: relation between delay discounting and health behavior. Journal of Behavioral Medicine. 2013; 36:34–43. [PubMed: 22311103]

Metcalfe J, Mischel W. A hot/cool-system analysis of delay of gratification: dynamics of willpower. Psychological Review. 1999; 106:3–19. [PubMed: 10197361]

Mischel W, Shoda Y, Rodriguez MI. Delay of gratification in children. Science. 1989; 244:933–938. [PubMed: 2658056]

Mitchell SH. The genetic basis of delay discounting and its genetic relationship to alcohol dependence. Behavioural Processes. 2011; 87:10–17. [PubMed: 21354276]

Muthen, LK.; Muthen, BO. Mplus users guide. 5th edition. Los Angeles, CA: Muthen & Muthen; 1998–2008.

Myerson J, Green L, Warusawitharana M. Area under the curve as a measure of discounting. Journal of the Experimental Analysis of Behavior. 2001; 76:235–243. [PubMed: 11599641]

Navarick DJ. Discounting of delayed reinforcers: Measurement by questionnaires versus operant choice procedures. The Psychological Record. 2004; 54:85–94.

Neale, MC.; Cardon, LR. Methodology for genetic studies of twins and families. Dordrecht, Netherlands: Kluwer Academic Publisher; 1992.

Olson EA, Hooper CJ, Collins P, Luciana M. Adolescents' performance on delay and probability discounting tasks: contributions of age, intelligence, executive functioning, and self-reported externalizing behavior. Personality and Individual Differences. 2007; 43:1886–1897. [PubMed: 18978926]

Paloyelis Y, Asherson P, Mehta MA, Faraone SV, Kuntsi J. DAT1 and COMT effects on delay discounting and trait impulsivity in male adolescents with attention deficit/hyperactivity disorder and healthy controls. Neuropsychopharmacology. 2010; 35:2414–2426. [PubMed: 20736997]

Reynolds B, Fields S. Delay discounting by adolescents experimenting with cigarette smoking. Addiction. 2012; 107:417–424. [PubMed: 21906199]

Reynolds B, Ortengren A, Richards JB, De Wit . Dimensions of impulsive behavior: personality and behavioral measures. Personality and Individual Differences. 2006; 40:305–315.

Reynolds B, Penfold RB, Patak M. Dimensions of impulsive behavior in adolescents: laboratory behavioral assessments. Experimental and Clinical Psychopharmacology. 2008; 16:124–131. [PubMed: 18489016]

Reynolds B, Schiffbauer R. Measuring state changes in human delay discounting: an experiential discounting task. Behavioural Processes. 2004; 67:343–356. [PubMed: 15518985]

Reynolds B, Schiffbauer R. Delay of Gratification and Delay Discounting: A unifying feedback model of delay-related impulsive behavior. Psychological Record. 2005; 55:439–460.

Richards JB, Zhang L, Mitchell SH, de Wit . Delay or probability discounting in a model of impulsive behavior: effect of alcohol. Journal of the Experimental Analysis of Behavior. 1999; 71:121–143. [PubMed: 10220927]

Robins, LN.; Baber, T.; Cottler, LB. Composite International Diagnostic Interview: Expanded Substance Abuse Module. St. Louis, MO: Author [L.N.R.]; 1987.

Romer D, Duckworth AL, Sznitman S, Park S. Can adolescents learn self-control? Delay of gratification in the development of control over risk taking. Prevention Science. 2010; 11:319–330. [PubMed: 20306298]

Sattler, J. Assessment of children: Cognitive applications. San Diego: Jerome Sattler Publisher, Inc.; 2001.

Scheres A, Lee A, Sumiya M. Temporal reward discounting and ADHD: task and symptom specific effects. Journal of Neural Transmission. 2008; 115:221–226. [PubMed: 17876680]

Scheres A, Sumiya M, Thoeny AL. Studying the relation between temporal reward discounting tasks used in populations with ADHD: a factor analysis. Int J Methods Psychiatr Res. 2010; 19:167–176. [PubMed: 20812292]

Shamosh NA, Gray JR. Delay discounting and intelligence: A meta-analysis. Intelligence. 2008; 36:289–305.

Smith CT, Boettiger CA. Age modulates the effect of COMT genotype on delay discounting behavior. Psychopharmacology. 2012; 222:609–617. [PubMed: 22349272]

Sparks J, Isen J, Iacono W. Preference on cash-choice task predicts externalizing outcomes in 17-year-olds. Behavior Genetics. 2014; 44:102–112. [PubMed: 24442381]

Steinberg L, Graham S, O'Brien L, Woolard J, Cauffman E, Banich M. Age differences in future orientation and delay discounting. Child Development. 2009; 80:28–44. [PubMed: 19236391]

Tellegen, A.; Waller, NG. Exploring personality through test construction: Development of the Multidimensional Personality Questionnaire. In: Boyle, GJ.; Matthews, G.; Saklofske, DH., editors. Handbook of personality theory and testing: Personality measurement and assessment. Vol. II. London: Sage; 2008. p. 261-292.

Vaidya JG, Latzman RD, Markon KE, Watson D. Age differences on measures of disinhibition during young adulthood. Personality and Individual Differences. 2010; 48:815–820.

Wilson VB, Mitchell SH, Musser ED, Schmitt CF, Nigg JT. Delay discounting of reward in ADHD: application in young children. Journal of Child Psychology and Psychiatry, and Allied Disciplines. 2011; 52:256–264.

Wulfert E, Block JA, Santa Ana E, Rodriguez ML, Colsman M. Delay of gratification: impulsive choices and problem behaviors in early and late adolescence. Journal of Personality. 2002; 70:533–552. [PubMed: 12095190]

NIH-PA Author Manuscript
NIH-PA Author Manuscript

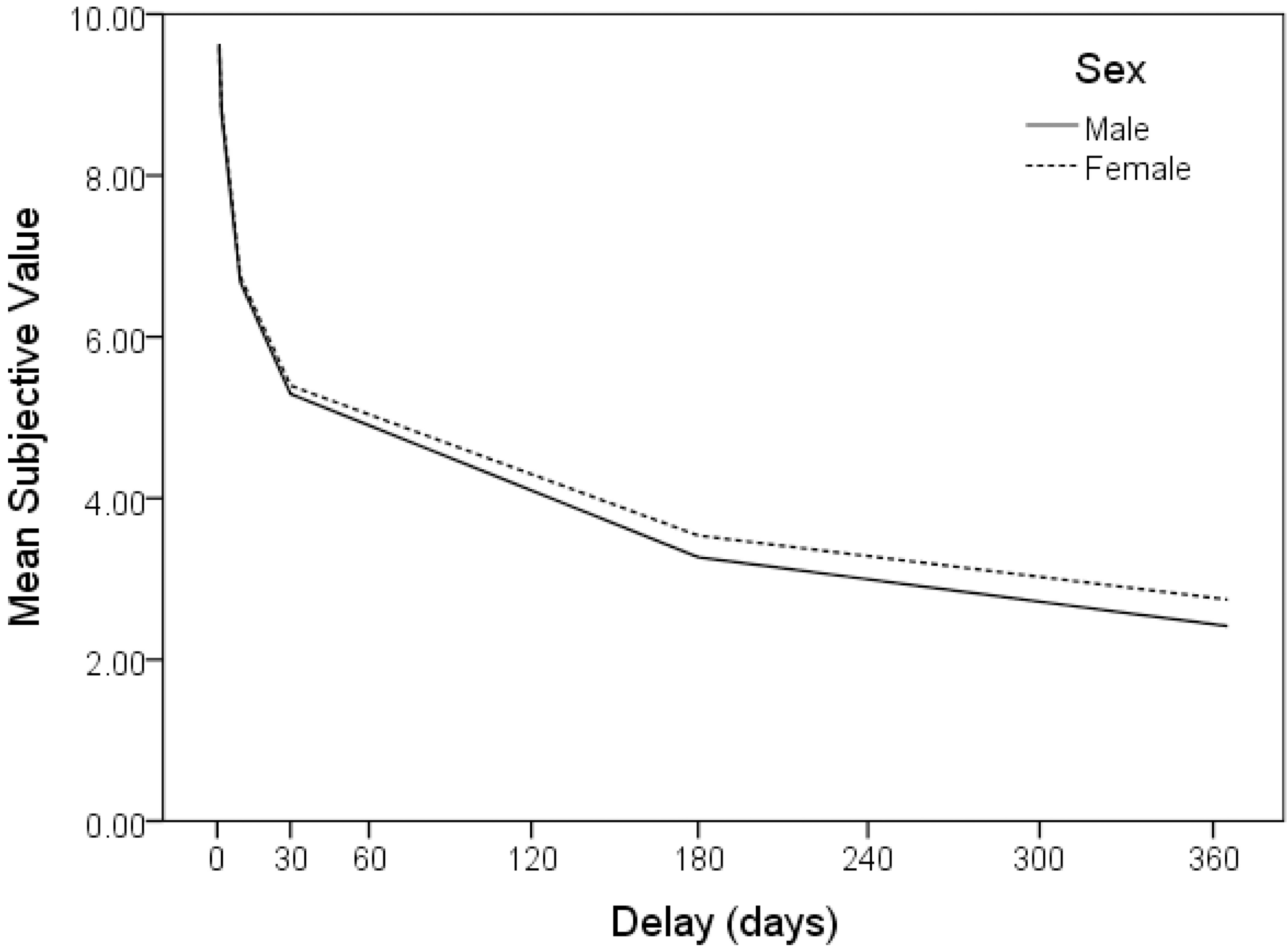

Figure 1.
Distribution of indifference points in males and females. Each point corresponds to the mean indifference amount at a given delay.

Table 1

Descriptive Statistics in Males and Females

Measure	Males		Females	
	Mean (*SE*)	***N***	**Mean (*SE*)**	***N***
AUC	−0.03 (0.06)	303	0.02 (0.07)	254
Cash Choice	0.73 (0.03)	392	0.80 (0.02)	399
Vocabulary	10.55 (0.15)	390	10.38 (0.17)	402
Block Design	12.00 (0.18)	388	11.70 (0.17)	402
SES	0.07 (.07)	478	−0.07 (0.06)	520
Early SU	0.28 (0.03)	391	0.27 (0.03)	399
EXT[a]	0.19 (0.05)	442	−0.18 (0.06)	471
MPQ Control[a]	−0.08 (0.05)	453	0.06 (0.05)	502

Notes. AUC = area-under-the-curve (Blom transformed); Cash Choice = proportion of participants selecting the delayed reward; Vocabulary and Block Design are scaled-score equivalents from the WAIS-R; SES = socioeconomic status; Early SU = proportion of participants with nicotine, cannabis, or alcohol use experience by age 14–15; EXT = externalizing liability; MPQ Control is a scale from the Multidimensional Personality Questionnaire.

[a]mean sex difference, $p < .05$

Table 2

Polychoric and Polyserial Correlations among Cash Choice Outcome, AUC, and Indifference Amounts

	Cash Choice	**AUC**
Cash Choice	---	
AUC	.23	---
1-day	.13[a]	.65
2-day	.11[a]	.71
10-day	.21	.87
30-day	.17	.94
180-day	.28	.96
365-day	.27	.93

Notes. All variables are categorical except for AUC. Indifference amounts and AUC are measured from the same task.

[a]polychoric correlation is not significant, $p > .05$

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 3

Correlation Matrix for Males and Females

	AUC	Cash Choice	Vocab	Block	SES	Early SU	EXT	MPQ Control
AUC	---	.22	.21	.28	.14	.06[a]	−.05[a]	.00[a]
Cash Choice	.24	---	.36	.31	.21	−.35	−.38	.26
Vocabulary	.17	.25	---	.31	.43	−.30	−.27	.18
Block Design	.23	.22	.29	---	.20	−.10[a]	−.03[a]	.11
SES	.23	.24	.32	.16	---	−.32	−.26	.16
Early SU	−.23	−.31	−.25	−.21	−.16	---	.72	−.42
EXT	−.10[a]	−.44	−.23	−.16	−.13	.74	---	−.50
MPQ Control	.11[a]	.20	.25	.05[a]	.10	−.40	−.43	---

Notes. Males are listed in the lower triangular portion; females in the above portion. All values refer to Pearson product-moment or biserial correlation coefficients, excepting the association between cash choice and early SU (which is tetrachoric).

[a]correlation is not significant, $p > .05$