Behavioural Processes 111 (2015) 55–59

Contents lists available at ScienceDirect

Behavioural Processes

journal homepage: www.elsevier.com/locate/behavproc

Short report

Long-term test–retest reliability of delayed reward discounting in adolescents

Andrey P. Anokhin*, Simon Golosheykin, Richard C. Mulligan

Washington University School of Medicine, St. Louis, MO, USA

ARTICLE INFO

Article history:
Received 2 April 2014
Received in revised form 3 October 2014
Accepted 10 November 2014
Available online 15 November 2014

Keywords:
Decision making
Delay discounting
Impulsivity
Reward

ABSTRACT

Delay discounting (DD), a decline in subjective value of a reward with increasing temporal delay in receipt of that reward, is an established behavioral indicator of impulsivity. Preference for smaller-immediate over larger-delayed rewards has been implicated in the basic neurobehavioral mechanisms of risk for addictive disorders and related externalizing psychopathology. Establishing long-term stability of DD in adolescence is a necessary step towards its validation as an intermediate phenotype, or marker of risk, in neurobiological and genetic studies. Previous studies have demonstrated moderate to high test–retest reliability of DD, however, these studies utilized adult samples and examined relatively short retest intervals. Due to continuing development of brain and behavior, stability of temporal discounting behavior in adolescence may differ from that in adulthood. Here, two cohorts of adolescents aged 16 ($n = 126$) and 18 ($n = 111$) were administered a computerized test of DD and re-tested two years later. DD rate showed a modest but significant decrease with age, suggesting a reduction in overall impulsivity from middle to late adolescence. Significant test–retest correlations were observed in both cohorts (.67 and .76, respectively, $p < .001$) indicating longitudinal stability of individual differences in decision-making behavior during middle and late adolescence.

This article is part of a Special Issue entitled: insert SI title.

1. Introduction

Impulsive choice is a distinct component of a broader, multifaceted “impulsivity” construct. It is commonly operationalized through delay discounting (DD) paradigms. Delay discounting refers to decrease in the subjective value of a reward with increasing temporal delay in receipt of that reward (Reynolds, 2006; Reynolds et al., 2006). Steeper DD reflects a tendency to choose smaller but immediate rewards over larger but delayed rewards. Human and animal studies have shown that DD is implicated in the basic biobehavioral mechanisms that underlie addictive behaviors and other externalizing psychopathology (reviewed in Bickel et al., 2013; Dalley et al., 2011; Mackillop, 2013; Reynolds, 2006).

Evidence that has linked risk for addiction to substances of abuse with the propensity for discounting delayed rewards suggests that DD may be an intermediate phenotype (endophenotype) for a range of disorders characterized by high levels of impulsivity, including substance use disorders (SUD) as well as attention deficit hyperactivity disorder (ADHD) and conduct disorder (CD). A focus on the relatively homogenous component processes comprising liability to psychiatric disorder may be more fruitful than examining the complex diagnostic phenotypes themselves and might facilitate identification of the neurobiological and genetic underpinnings of addiction and psychopathology.

An important prerequisite for such an intermediate phenotype is its intra-individual stability. Previous research has shown that DD measures represent a stable, trait-like characteristic. Across several studies, estimates of the test–retest reliability of DD measures ranged from .55 to .90 (Baker et al., 2003; Beck and Triplett, 2009; Johnson et al., 2007; Kirby, 2009; Ohmura et al., 2006; Simpson and Vuchinich, 2000; Smits et al., 2013; Weafer et al., 2013). However, several aspects of these previous studies warrant further investigation.

First, previous studies were based on data collected from adult participants. Recently, there has been increasing interest in DD as an experimental measure of impulsivity during adolescence, a period of development marked by increased risk for impulsive behaviors and substance abuse (Bava and Tapert, 2010; Crews et al., 2007). However, due to continuing brain development which primarily includes the regions that are critically involved in decision making (Casey et al., 2005; Paus, 2005), this period is characterized by significant cognitive and behavioral changes. Consequently,

* Corresponding author. Tel.: +1 314 286 2201; fax: +1 314 286 0091.
E-mail address: andrey@wustl.edu (A.P. Anokhin).

http://dx.doi.org/10.1016/j.beproc.2014.11.008
0376-6357/

the stability of DD behavior in adolescence may differ from that in adulthood. Therefore, data obtained in adult samples cannot be directly generalized to the adolescent population, and it is important to determine the long-term stability of DD during this period of major developmental change.

Second, most previous studies used short retest intervals (typically, a few weeks), with the notable exception of one study (Kirby, 2009) that used a retest interval of one year and showed significant test–retest stability of DD assessed using a 27-item questionnaire. Third, most laboratory studies of DD have been based on small samples ($n = 15–33$) making it difficult to precisely estimate test–retest reliability. For example, with a sample size of $n = 30$, the 95% confidence interval of a correlation of $r = 0.6$ would range between 0.31 and 0.79 (i.e., from "low" to "high"). One notable exception is a recent study by Weafer et al. (2013) that included a large sample of 128 young adults and found a high test–retest correlation for DD ($r = 0.89$), however, mean retest interval was only about 9 days.

One of our recent studies demonstrated significant longitudinal stability and heritability of inter-temporal choice in adolescents using a real-money, single-trial delayed gratification procedure (Anokhin et al., 2011). However, little is known about developmental test–retest stability of more commonly used indices of DD that are based on varying amounts of hypothetical money and differing delays in receipt of those rewards. The purpose of the present study was to assess longitudinal stability of individual differences of DD measures in middle to late adolescence using a well-powered, population-representative sample and a version of the DD paradigm that can probe how manipulation of reward amount and temporal delays to receipt of those rewards can affect subjective value of those rewards and decision-making. An additional aim was to determine whether DD undergoes systematic changes during middle and late adolescence.

2. Material and methods

2.1. Participants

Two longitudinal cohorts of adolescents participated in the study: 16-years-olds ($n = 126$, 65 females, M age ± SD: 16.6 ± .26) and 18-years-olds ($n = 111$, 59 females, M age ± SD: 18.7 ± .27). The sample was 84% Caucasian, 12% Black, and 4% were other minorities. Participants were administered a computerized DD test (described below) twice with an average interval between the two test administrations of approximately 2 years (age 16 cohort: 24.7 ± 2.2 months; age 18 cohort: 23.5 ± 2.2 months). Thus, DD data were collected at ages 16 and 18 for the younger cohort and at ages 18 and 20 for the older cohort, however, no data were available yet to compare DD at ages 16 and 20. All individuals included in the present analysis participated in a larger ongoing twin study of brain, cognition, and behavior. They were originally ascertained randomly through use of a state birth records database to ensure that the sample was maximally representative of the general population. Therefore, the present sample matched well the general population with respect to variables that potentially could bias the results, such as general intelligence (as assessed using Raven's Standard Progressive Matrices test) and socioeconomic status (parents' report). Exclusion criteria were minimal but included inability to understand task instructions, uncorrectable vision or hearing impairments, or current illicit drug or alcohol intoxication (verified by breathalyzer and urine drug test). Participants with reporting recent (one week) drug use or current intoxication or showing positive results on the urine drug test were excluded. Additionally, participants showing positive results on the alcohol breathalyzer test were excluded. Tobacco smokers were given an opportunity to smoke when they arrived for the laboratory session to minimize potential nicotine withdrawal effects, but not within the last 60 min before the start of assessments. All experiments on human subjects were conducted in accordance with the declaration of Helsinki. The study was approved by Washington University Institutional Review Board and written informed consent was obtained from all participants.

2.2. Discounting task

We used a computerized delay discounting task described previously (Mitchell, 1999; Mitchell and Wilson, 2010). In this task, participants were presented with a series of hypothetical choices wherein participants chose between a variable hypothetical amount of money available immediately and a delayed "standard" amount of $100 presented at one of six possible temporal delays. Questions and response options were presented on a computer screen, and participants used the computer mouse to make their responses. On each of the 138 trials of this task, participants were presented with a question: "At this moment, what would you prefer?" Two choice options were displayed underneath this question (e.g. "$100 in 90 days", "$30 now"). One option was always a standard amount of $100 available after one of six delays: 0, 7, 30, 90, 180, or 365 days. The other option was an alternative amount of money available immediately. The test questions were formed by the combination of 6 standard amounts ($100 available at one of 6 delays) and 23 alternative amounts (ranging from $5 to %105 and available immediately, i.e. at 0 delay), which resulted in a total of 138 questions. One question that would require participants to choose between identical items ($100 now or $100 in 0 days) was omitted, thus the final set included 137 choice questions. For each question, a pair of immediate and delayed amounts was selected at random without replacement. The order in which the immediate and delayed amount was presented (first or second in the pair) was varied randomly. Participants made their choices by indicating the preferred option with a computer mouse. As response time was not limited, the duration of the task was variable but it was typically completed within 10 min. Further details of the procedure can be found elsewhere (Mitchell, 1999; Mitchell and Wilson, 2010).

2.3. Discounting measures

For each of the 6 delays of the standard amount ($100), an indifference (switch) point was determined. The indifference point was defined as being midway between the smallest value of the immediate alternative that was accepted and the largest value of the immediate alternative that was rejected (Mitchell and Wilson, 2010). That is, the indifference or switch point is the value at which the participant acted as if they were indifferent to the choice between a particular amount that was available immediately and the delayed standard amount. This amount may be viewed as the "subjective value" of the delayed standard amount for that particular individual when receipt of the standard amount is delayed by a specific period of time. Indifference points can be inferred from an individual's pattern of choices, i.e. switching from the immediate to the delayed reward and vice versa as a function of change in the amount of immediate reward. For example, if the participant's choice switched to a delayed $100 reward available in 180 days when immediate reward option dropped to less than $65 or, conversely, the choice switched to immediate reward when it exceeded $65, then, for a given individual and a given delay, $100 to be received in 180 days is subjectively worth only $65. Next, we built an empirical discounting curve by plotting indifference values against the corresponding delay values (Fig. 1) and computed area under the curve (AUC) as a quantitative measure of DD. Smaller AUC values indicate a steeper function and thus greater degree of temporal discounting, i.e. preference of smaller immediate rewards