Choice Impulsivity: Definitions, Measurement Issues, and Clinical Implications

Kristen R. Hamilton1*, Marci R. Mitchell2*, Victoria C. Wing3*, Iris M. Balodis2, Warren Bickel1, Mark Fillmore5, Scott D. Lane6, C. W. Lejuez1, Andrew K. Littlefield7, Maartje Luijten8, Charles W. Mathias6, Suzanne H. Mitchell9, T. Celeste Napier10, Brady Reynolds11, Christian G. Schütz12, Barry Setlow13, Kenneth J. Sher7, Alan C. Swann14, Stephanie E. Tedford15, Melanie J. White16, Catharine Winstanley17, Richard Yi1, Marc N. Potenza2,18^, F. Gerard Moeller19^  

* indicates that these authors contributed equally to the generation of this manuscript  
^ indicates that these authors contributed equally to the generation of this manuscript

Corresponding Author:  
Department of Psychology  
University of Maryland  
2103 Cole Field House  
College Park, MD 20742  
Email: khamilt4@umd.edu

1Department of Psychology, Center for Addictions, Personality, and Emotion Research, University of Maryland, College Park, MD, USA 20742  
2Department of Psychiatry, Yale University School of Medicine, New Haven, CT, USA 06519  
6Department of Psychiatry, University of Texas Health Science Center San Antonio, San Antonio, TX 78229  
7 Department of Psychology, Texas Tech University, Lubbock, TX, USA 79409  
8 Behavioural Science Institute, Radboud University Nijmegen, Nijmegen, the Netherlands  
9Departments of Behavioral Neuroscience and Psychiatry, Oregon Health & Science University, Portland, OR 97239  
10Departments of Pharmacology and Psychiatry, Center for Compulsive Behavior and Addiction, Rush University Medical Center, Chicago, IL, USA 60612  
11Department of Behavioral Science, University of Kentucky, Lexington, KY, USA 40536  
15Department of Pharmacology, Center for Compulsive Behavior and Addiction, Rush University Medical Center, Chicago, IL, USA 60612  
16School of Psychology and Counselling, Institute of Health and Biomedical Innovation, Queensland University of Technology, Kelvin Grove, QLD, Australia 4059  
17Department of Psychology, University of British Columbia, Vancouver, British Columbia, Canada  
18Departments of Neurobiology and Child Study Center, Yale University School of Medicine, New Haven, CT, USA 06519  
19Departments of Psychiatry and Pharmacology and Toxicology, Virginia Commonwealth University School of Medicine, Richmond, VA, USA 23219
Abstract

**Background:** Impulsivity critically relates to many psychiatric disorders. Given the multifaceted construct that impulsivity represents, defining core aspects of impulsivity is vital for the assessment and understanding of clinical conditions. Choice impulsivity (CI), involving the preferential selection of smaller sooner rewards over larger later rewards, represents one important type of impulsivity. **Method:** The International Society for Research on Impulsivity (InSRI) convened to discuss the definition and assessment of CI and provide recommendations regarding measurement across species. **Results:** Commonly used preclinical and clinical CI behavioral tasks are described, and considerations for each task are provided to guide CI task selection. Differences in assessment of CI (self-report, behavioral) and calculating CI indices (e.g., area-under-the-curve, indifference point, steepness of discounting curve) are discussed along with properties of specific behavioral tasks used in preclinical and clinical settings. **Conclusions:** The InSRI group recommends inclusion of measures of CI in human studies examining impulsivity. Animal studies examining impulsivity should also include assessments of CI and these measures should be harmonized in accordance with human studies of the disorders being modeled in the preclinical investigations. The choice of specific CI measures to be included should be based on the goals of the study and existing preclinical and clinical literature using established CI measures.

**Key Words:** impulsivity, response, attention, behavioral control, personality, delay discounting, delay of gratification, self-control
List of Abbreviations: CI: Choice impulsivity; DD: delay discounting; TD: temporal discounting; IC: intertemporal choice; ITCT: inter-temporal choice tasks; PD: personality disorders; InSRI: International Society for Research on Impulsivity; EDT: Experiential Discounting Task; MCQ: Monetary Choice Questionnaire.
**Introduction**

Among its many definitions, impulsivity has been described as a “predisposition toward rapid, unplanned reactions to internal or external stimuli with diminished regard to the negative consequences of these reactions to the impulsive individual or to others” (Evenden, 1999; Moeller, et al., 2001; Potenza & de Wit, 2010). Impulsivity has been related to multiple psychiatric conditions including bipolar, substance-use and many different personality disorders (Chamorro, et al., 2012; Moeller, et al., 2001). Given that measures of impulsivity have been linked to clinically relevant constructs like treatment outcomes (Blanco, et al., 2009; Krishnan-Sarin, et al., 2007; MacKillop & Kahler, 2009; Stanger, et al., 2012), precise and consistent assessments of impulsivity may help improve clinical care provided to multiple psychiatric populations.

There has been considerable discussion and some debate regarding impulsivity’s precise boundaries and components (Gullo, et al., 2014). For example, a recent review of impulsivity and its associated neurobiological correlates described facets relating to response, choice, reflection and decision-making (Fineberg, et al., 2014), and some researchers have advocated for a parsimonious approach to considering the main factors or domains contributing to impulsivity (Gullo, et al., 2014). Based on the definition of impulsivity above, lack of planning and lack of regard for future consequences are both important features of impulsivity, as is rapid responding to external and internal stimuli. Consistent with the multiple dimensions of impulsivity based on the definition, prior studies have shown that impulsivity is a multi-faceted construct of which two or more components have typically been identified in factor analyses of both self-report and behavioral measures (Broos, et al., 2012; Meda, et al., 2009; Reynolds, et al., 2008; Verdejo-
Garcia, et al., 2008). One component of impulsivity, rapid response impulsivity (discussed in the accompanying manuscript), may be further fractionated into impulsive actions relating to refraining from initiating an action versus stopping an action that has been initiated (see Fineberg et al, 2014 and accompanying manuscript). A second component, choice impulsivity (CI), refers to making impulsive decisions and involves tendencies to select smaller-sooner rewards over larger-later rewards (e.g., the choosing of immediate but smaller versus delayed and larger rewards) and may relate to difficulties in delaying gratification or exerting self-control (Fineberg, et al., 2010). In a meta-analysis of impulsivity-related measures, there was moderate convergence between CI and trait impulsivity (as assessed by self-report and informant-report questionnaires) (Duckworth & Kern, 2011).

CI includes two aspects included in the definition of impulsivity, lack of planning and lack of regard for future consequences (Blakemore & Robbins, 2012; Grant & Chamberlain, 2014; Hamilton & Potenza, 2012; Peters & Buchel, 2011). Additional terms that have been used to describe this component include delay discounting and temporal discounting, among others (Fineberg, et al., 2014; Fineberg, et al., 2010). Further evidence supporting CI as a dimension of impulsivity includes results from studies showing that groups with elevated impulsivity on a clinical level also have higher CI, and these groups include individuals with borderline personality disorder, bipolar disorder, and addictions (Ahn, et al., 2011; Lawrence, et al., 2010). CI also has been suggested to be a trans-diagnostic process underlying addictions, gambling, obesity, poor health practices, and financial mismanagement underscoring the importance of delay-discounting research to public health (Bickel, et al., 2012b; Hamilton & Potenza, 2012). The need for studying trans-diagnostic processes recently has received increased emphasis (e.g., by the RDoC initiative of the NIMH) (Insel, et al., 2010)).
The extent to which other processes relating to information gathering (as assessed by information-sampling tasks) (DeVito, et al., 2009) and risk/reward decision-making more generally (as assessed by such tasks as the Iowa Gambling Task) (Bechara, et al., 1994) are encompassed by CI or distinct from CI has been debated and is an area of active research (Broos, et al., 2012; Fineberg, et al., 2014; Fineberg, et al., in press; Verdejo-Garcia, et al., 2008). In this article, we will consider CI as the preferential selection of smaller-sooner rewards over larger-later rewards, and use the term CI to refer to such a process, typically assessed behaviorally through the use of inter-temporal choice tasks (ITCTs).

As the term impulsivity has been used in multiple and varied fashion, the goal of the current manuscript is to review and make recommendations regarding the assessment of CI. The manuscript follows the 2013 meeting of the International Society for Research on Impulsivity (InSRI), which was devoted to the discussion of this topic, and a working group that resulted to review and synthesize information for inclusion in this manuscript. The manuscript reviews non-human (particularly rodent) and human research assessments of CI and the use of the latter in clinical populations. The manuscript concludes with recommendations based on existing findings.

**Choice Impulsivity Relates Prospectively to Harmful Behaviors**

Studying CI is critical because it relates prospectively to detrimental behaviors (Audrain-McGovern, et al., 2009; Chabris, et al., 2008; Fernie, et al., 2013; Kishinevsky, et al., 2012; Yoon, 2007) and is reliably elevated in multiple, relevant patient populations (Caceda, et al., 2014; Rogers, et al., 2010). Additionally, the translational aspect of the CI construct facilitates its measurement in animals and humans.
CI is associated with a variety of problematic behaviors, including gambling (Leeman & Potenza, 2012; Madden, et al., 2011), binge eating (Davis, et al., 2010), suicide (Dombrovski, et al., 2011), violence (Cherek & Lane, 1999; Cherek, et al., 1997), substance use (Kollins, 2003), and risky sexual behaviors (Johnson & Bruner, 2011). CI studies have revealed higher levels of delay discounting in abusers of many drugs, including alcohol (Lejuez, et al., 2010; Petry, 2001), tobacco (Reynolds, et al., 2007; Reynolds, et al., 2004; Sweitzer, et al., 2008), cocaine (Bornovalova, et al., 2005; Heil, et al., 2006; Kirby, 2004; Petry, 2003), and heroin (Kirby, 2004; Petry, 2003), than by non-abusing control subjects. CI is associated with overall substance-use patterns (e.g. Lejuez et al., 2010) and substance use on a momentary basis. For example, within a single drinking session, baseline levels of CI relate prospectively to blood alcohol level in social drinkers when they consume alcohol (Moore & Cusens, 2010). Not only does substance use occur more often in individuals with high levels of CI, substance use itself may increase CI (Bickel, et al., 1999; Petry, 2001; Sweitzer, et al., 2008).

CI also impacts sexual decision-making. Impulsive choices about hypothetical sexual outcomes and erotic stimuli on delay discounting tasks (Lawyer, 2008; Lawyer, et al., 2010) conform to the hyperbolic-like function that describes the discounting of non-sexual delayed consequences (Myerson & Green, 1995; Rachlin, et al., 1991). The contribution of CI to risky sexual behavior was further evident in a hypothetical Sexual Discounting Task (Johnson & Bruner, 2011) in which there was an effect in which participants chose to engage in immediate unprotected sex rather than use a condom when obtaining a condom involved a delay. Further, the effect of CI to decrease condom use is this effect was correlated with self-reported real-life risky sexual behaviors, underscoring the importance of CI to public health (Johnson & Bruner, 2011).
The important role of CI in health-risk behaviors is evidenced by longitudinal research in which CI levels were prospectively related to subsequent health risk behaviors, including smoking acquisition in alcohol involvement in adolescents (Audrain-McGovern, et al., 2009; Chabris, et al., 2008; Fernie, et al., 2013; Kishinevsky, et al., 2012; Yoon, 2007). Not only does CI prospectively relate to the acquisition of health risk behaviors, it also prospectively relates to substance abuse treatment outcomes in adolescents and adults (MacKillop & Kahler, 2009; Sheffer, et al., 2013; Stanger, et al., 2012). The predictive relationship between CI and substance use outcomes observed in clinical research also is evident in preclinical research. For example, in a rodent model of nicotine addiction, CI prospectively related to nicotine seeking during abstinence and an enhanced vulnerability to relapse when exposed to nicotine cues (Diergaarde, et al., 2008). Similarly, in a rodent model of cocaine addiction, CI related prospectively to extinction and propensity to relapse in the context of cocaine cues (Broos et al., 2012). Clinical and preclinical research evidencing the predictive validity of CI in health risk behaviors speaks to the importance of including measures of CI in research.

**Assessing Choice Impulsivity**

ITCTs involve choices between smaller-sooner rewards and larger-later rewards. A series of choices are presented in order to determine the extent to which individual preferences for smaller-sooner versus larger-later rewards exist. Greater discounting of the large reward (e.g., steeper discounting) is indicative of more impulsive behavior, although differences in temporal judgment and reward sensitivity may likewise contribute to such propensities. Similarly, as delayed rewards carry variable degrees of uncertainty in real life, some ITCTs incorporate a probabilistic element. Thus, while sensitivities to delay and uncertainty may influence decisions
to select options of differing reward magnitudes and individuals may show varying degrees of probabilistic discounting, the current manuscript will focus on predominantly on assessing delay discounting.

Several measures of impulsive choice may be derived from ITCTs including an indifference point between the two options, the area under the curve (AUC), and percent choice of the large (or small) reward. Indifference points refer to the change in preference from later to sooner or sooner to later, and can occur at any percentage where preferences change.

The indifference points for a series of delays and rewards can be used to produce a discount curve, which typically describes the rate at which the value of a reward decreases as a function of increased time to reward receipt (Mazur, 1987; Mendez, et al., 2010; Odum, 2011; Richards, et al., 1997). Moreover, this method permits investigations of subjective ‘easy’ versus ‘hard’ choices by those trials that deviate more from an individual’s indifference point (Amlung, et al., 2012; Hoffman, et al., 2008; Monterosso, 2007). Indifference points for a series of reward/delay-time choice options can be used to create a discount curve to reflect the rate at which the value of a reward decreases as a function of increased time to reward receipt.

Characterizing patterns of delay discounting is useful for helping investigators to more fully understand CI and make better predictions regarding CI-related behavioral outcomes. In a hyperbolic discounting function, the value of a reward declines rapidly for small delay periods but more slowly for longer delay periods. By contrast, in exponential discounting the reward value is discounted by a factor that increases with the length of delay. Hyperbolic discounting is the most commonly used mathematical model, and research suggests that discounting follows hyperbolic, rather than exponential, trajectories (Mazur & Biondi, 2009; Rachlin, et al., 1991).
The hyperbolic discounting function can be summarized by the equation $V = A / (1 + kD)$, in which $V$, the value of the delayed reinforcer (present value of the reward or indifference point), is equal to the amount of the reinforcer ($A$), divided by the delay to the reward ($kD$). In this equation, $k$ is a free parameter that describes the steepness of the discount function (i.e., the scaling factor which manipulates $kD$ and describes the degree to which value is affected by the delay) (Dallery & Locey, 2005; Mazur, 1987; Odum, 2011; Richards, et al., 1997). A higher $k$ reflects greater CI. The benefit of using a $k$ value as the index measure of delay discounting is that it is relatively stable and has test-retest reliability. For example, Kirby and colleagues found that with repeated testing using similar test situations, individuals had similar discounting rates (i.e., $k$ values) up to one year later (i.e., test-retest reliability = .71) on a questionnaire measure of discounting (Kirby, 2009). A potential downfall of fitting a curve using this mathematical model is that while the overall fit may be good, it may overestimate indifference points when delays are short and overestimate points when delays are longer.

By contrast, the AUC method may more directly represent patterns of indifference points. To calculate the AUC, the delay and associated subjective value (or response on the large reinforcer lever in animal ITCTs) are expressed as a proportion of their respective maximum values. The normalized values are used as $x$- and $y$-coordinates, respectively, and the AUC is calculated by summing the results of $(x_2-x_1)(y_1+y_2)/2$, where $x_1$ and $x_2$ are successive delays and $y_1$ and $y_2$ are the subjective values associated with those delays. AUC values range between 0.0 (steepest possible discounting) and 1.0 (no discounting), larger AUCs represent less discounting or lower CI (Myerson, et al., 2001; Odum, 2011). As with indifference points and percent choice, AUCs are not dependent upon theoretical assumptions regarding the form of the discounting functions (Myerson, et al., 2001; Odum, 2011). However, this feature also constitutes a
disadvantage as the data represent the subjective value expressed as a proportion of the nominal value and thus AUCs from different experiments may not be appropriately compared without adjusting for differences in range. A second disadvantage is that the AUCs of two discounting functions may be the same even though the two functions have different shapes (Myerson, et al., 2001). However, neither AUC’s nor $k$ values represent a direct measure of behavior. A third disadvantage is that in a small proportion of cases indifference points for individual subjects cannot be fitted to the hyperbolic equation due to irregular choice patterns.

In preclinical research, another measure of discounting is percent choice, which typically refers to the percent choice of the large, delayed reinforcer at different delays as the dependent measure. A benefit of this measure is that it directly indexes the observed behavior. This method may be used when the responses to multiple delays are assessed within a single session (“within-session shifts” design) or in multiple sessions (“between-session shifts” design) (Evenden & Ryan, 1996; Mitchell & Wilson, 2012; Winstanley, et al., 2003). Some benefits of percent choice include ease of calculation and relative insensitivity to the number of responses or omitted trials. However, it is of note that the number of omissions can vary depending on the method used to analyze percent choice (e.g., percent selections out of either total trials or total responses), which may require researchers to set a criterion for an acceptable number of omissions. A caveat of percent choice, however, is that it complicates collapsing choice behavior over different delays into a single value as it confounds discounting rate at specific delays with overall responding. For example, 5 different delay contingencies are often used in within-session designs. As with AUC, choice behavior between individual animals may result in very different curves but yield the same ‘overall’ or collapsed value for percent choice. As a result, most studies of this kind do not yield a single “percent choice” value by which subjects can be compared, but rather represent
data as a series of points, with percent choice at different delays representing important comparators. However, this is not as ideal as a single value by which individuals may be classified.

In summary, there are several different measures of impulsive choice that can be obtained from ITCTs. Each of the different measures (k, AUC, percent choice etc.) has advantages and disadvantages. To our knowledge, the predictive validity of the different measures of CI has not been compared empirically, which represents an important direction for future research. In addition, the reporting of multiple measures of CI could facilitate comparisons across studies.

*The Neurocircuitry of Choice Impulsivity*

CI can be conceptualized as the manifestation of an imbalance between neurobiological systems subserving control and motivation. In CI, choices for a small but immediate outcome are associated with activation in reward-related areas including the ventral striatum (VS) and medial prefrontal cortex (mPFC), while choices for a larger delayed outcome have been associated with activation in cortical areas including the dorsolateral (dl) and ventrolateral prefrontal cortex (vLPFC) (McClure et al., 2004). Alternate conceptualizations exist in which a distinct segregation is less clear (Kable & Glimcher, 2007).

Experiments involving the manipulation of PFC activation provide evidence supporting regional contributions. Temporarily increasing PFC activation directly resulted in reduced CI during the manipulation (Sheffer, et al., 2013), while temporarily decreasing PFC activation resulted in increased CI (Figner, et al., 2010). PFC hypoactivation underlying delay discounting also prospectively relates to future health risk behaviors that are associated with CI. In obese women, lower prefrontal and parietal activation during a delay discounting task was associated
with future weight gain (Kishinevsky, et al., 2012). Consonant with these findings, PFC activation prospectively related to successful weight loss in obese subjects (Weygandt, et al., 2013).

Connectivity of the hippocampus (HC) with the medial rostral prefrontal cortex (mrPFC) and anterior cingulate cortex (ACC) appears important in reducing delay discounting (Benoit, et al., 2011). Imagining future rewards attenuated delay discounting, and this effect was related to connections of the HC to the mrPFC (Benoit, et al., 2011) and ACC (Peters & Buchel, 2011). The HC has been implicated in scene construction (Hassabis & Maguire, 2007), and the effect of imagining future rewards to attenuate CI may be based on an interaction between regions supporting the construction of events (i.e., HC) and the representation of the event’s reward magnitude (i.e., mrPFC; Benoit et al., 2011).

**Animal Models of Choice Impulsivity**

While there are several studies examining nonhuman primates (Woolverton, et al., 2007), the majority of studies examining CI use rodents performing ITCTs that are similar in design to those used in humans. In these tasks, rodents are placed in operant chambers (although T-mazes can also be used; (Mariano, et al., 2009; Papale, et al., 2012)) and are presented with choice opportunities between smaller-sooner or larger-later rewards (or more precisely reinforcers, although we will use reward throughout to maintain consistency with the human literature discussed elsewhere in the manuscript; the reward in these studies is usually food or water, although intra-cranial self-stimulation can also be employed (Rokosik & Napier, 2011)). An animal’s CI is indexed by the time delay to receipt of reward at which the animal prefers the smaller-sooner reward. Rodent measures of delay discounting have demonstrated a high-level of test-retest reliability, with individual equivalence point scores remaining highly consistent over
time when tested for the first time in early adolescence (postnatal day 28-42) and then re-tested in adulthood (postnatal day 58-64) (McClure, et al., 2014).

Several systematic manipulations or variations can be implemented in animal ITCTs. ITCTs in animals may be conducted with either within-session or between-session shifts in delays. In a within-session design, each delay condition is presented to the animal in each test session, generally in blocks of trials using the same delays (for example, block 1 may contain a 0-sec delay to the large reward, block 2 a 10-sec delay, block 3 a 20-sec delay, etc.) (Evenden & Ryan, 1996; Mitchell & Wilson, 2012; Winstanley, et al., 2003). Each of these blocks usually begins with several forced choice trials in which only one lever is extended at a time, “forcing” the animal to experience the contingencies of each lever (i.e., which lever is designated the smaller-sooner reward and which is designated the larger-later reward). Following these ‘forced choice’ trials are ‘free choice’ trials on which both levers are extended and the animal can choose which lever they prefer.

In a between-session design, the delay associated with the larger-later reward does not shift within a test session, as previously described, but rather over the course of days. In this protocol, a single delay to the larger-later reward (e.g., 10 sec) is presented for the entire test session for one or several sessions before the delay is changed (Adriani & Laviola, 2003). To create a discounting function a series of delays are examined over the course of several sessions. There are several methods used in the context of between-sessions designs: infrequently percent choice of the larger delayed reward is examined (Poulos, et al., 1995), but more frequently one alternative is adjusted during the session to obtain an indifference point in a manner analogous to that used in several tasks used with humans. Basically, with this procedure, the animal’s choices during the session determine whether in subsequent trials the magnitude of the smaller-sooner
reward will be adjusted (adjusting amount procedures), or the delay to which the larger-later reward is received will be adjusted (adjusting delay procedure). For example, if the animal favors the larger-later reward, either the size of the smaller-sooner reward (adjusting amount procedure) or the delay to the larger-later reward (adjusting delay procedure) is increased. The measure of interest is the adjusting amount or delay at which animals are indifferent between the two choice alternatives.

Regardless of procedure, most studies include a condition/block for which there is a 0-s delay to the larger reward. The 0-second delay trial has several benefits. It can be used to assess potential alterations in reward-magnitude sensitivity driving changes in the delay discounting function (da Costa Araujo, et al., 2010). In addition, delay orders can be varied either between blocks (within-sessions) or between sessions to provide assurance that the animals have learned the task and that the responding profile does not reflect a side bias or perseveration (Tanno, et al., 2014). Additionally, the 0-sec delay may illuminate long-term carryover effects of repeated training in the task, such as a conditioned place aversion to the delayed response lever (Wilhelm & Mitchell, 2010). Such a delay condition may be particularly helpful when training mice. Whether a rodent has developed an aversion to the delayed side (i.e., the area around the delayed nose-poke aperture) can be tested by measuring the amount of time the rat spends in this area of the box. Murine versions of ITCTs are generally compressed in that they require smaller rewards, fewer trials, and shorter delays than do those used in rats (details of procedures for measurement of delay discounting in mice are discussed in (Mitchell, 2014)). Typically training mice is more onerous (with more intermediate steps), although whether this is due to issues associated with maintaining motivation or attention, the development of biases, or insensitivities to contingencies is unclear (Mitchell, 2014; Mitchell, et al., 2006). Reward magnitudes (e.g., 1
pellet each) or delays (e.g., 0-sec delay) may be equated between 2 choices, or the order in which the delays are presented may be reversed, to determine whether mice can switch between the two reward magnitudes and if they understand the delay contingencies.

One possible advantage of the within-session design over the between sessions procedures is that it permits the effects of pharmacological agents on choice behavior with different delay lengths to be assessed more rapidly, as performance under a range of delay durations may be assessed in a single test session. In contrast, the between-sessions design requires multiple test sessions to assess the effects of a single manipulation across delay durations (which could lead to sensitization or tolerance to the effects of the manipulation), and this may have advantages and disadvantages. The between-sessions design also has the possible advantage that the behavior may be more rapidly acquired than in the within-sessions design (Foscue, et al., 2012) A benefit of adjusting procedures is that they are more similar to the tasks used in research with human subjects, in which both delays and reward magnitudes shift within a session, allowing for calculations of $k$ values. Tasks with both adjusting and non-adjusting procedures are useful for assessing the effects of chronic manipulations (e.g., knock-outs or lesions) or trait-like aspects of impulsive choice (Helms, et al., 2006; Mendez, et al., 2010; Winstanley, et al., 2004). Given that certain approaches (e.g., optogenetic or genetic modifications) are feasible in research of rodents but not humans, CI research in animal models offers complementary insight that can inform the study of personality disorders and other psychiatric conditions.

In summary, future animal studies should take into account the disorders being modeled and extant data in the corresponding human conditions when choosing the type of ITCT and other specific aspects of CI measures as discussed above.
Human Inter-temporal Choice Tasks (ITCTs)

Recent characterizations of CI as a trans-diagnostic process underlying addiction, gambling, obesity, poor health practices, and financial mismanagement underscore the importance of CI research to public health (Bickel, et al., 2012b; Hamilton & Potenza, 2012). ITCTs are used to assess CI in humans (Madden & Johnson, 2010; Odum, 2011). Most ITCTs require participants to make a binary choice between a smaller reward that is delivered sooner and a larger reward that is delivered later (Peters & Buchel, 2011). Although different modes have been used to administer ITCT tasks (i.e., computer-based and paper and pencil), both administration modes generate important data that may be comparable across delivery modalities (Smith & Hantula, 2008). In both human and animal research, hyperbolic curves typically provide a better fit to delay discounting data than do exponential curves (Mazur & Biondi, 2009; Rachlin, et al., 1991). A discussion of different ITCTs, variants and outcome measures follows (For overview, see Table 1).

The Richards Task

The Richards computerized delay-discounting task is distinguished from traditional ITCTs by the use of an adjusting-amount protocol by which choices of smaller-sooner rewards over larger-later rewards cause subsequent choice alternatives to be dynamically modified until an indifference point is reached. Additionally, the task incorporates a probabilistic component to some of the delayed rewards. The task also includes realized contingencies that enhance ecological validity, as participants are informed that, depending upon their choices during the task, they may receive varying amounts of money at the conclusion of the experiment or
sometime thereafter (Richards, et al., 1999). Variants of this procedure have been developed and used by numerous other research groups to assess CI in distinct clinical groups, including individuals with attention-deficit/hyperactivity disorder (ADHD) (Paloyelis, et al., 2010), adolescent smokers (Fields, et al., 2011), and social drinkers (Reynolds, et al., 2006). The Richards task has a high level of test-retest reliability, with $r=.89$ when the interval between the first and second assessments was approximately 8 days (Weafer, et al., 2013). Additionally, unlike many other behavioral CI tasks, delay and probability discounting on the Richards task positively correlate with self-report measures of impulsivity, thereby suggesting construct validity of this task (Richards, et al., 1999). While it has been suggested that a limitation of the task might involve a relative insensitivity to experimental manipulations (e.g., methylphenidate administration, acute alcohol administration) when compared with other tasks measuring delay discounting in the same experiment, other research has reported effects of $d$-amphetamine on CI as assessed using the Richards task (de Wit, et al., 2002).

**The Experiential Discounting Task (EDT)**

The Experiential Discounting Task (EDT) (Reynolds & Schiffbauer, 2004) is a computerized real-time assessment of temporal discounting processes. Participants make choices between a delayed (0-sec, 7-sec, 14-sec, 28-sec) probabilistic (35% chance of occurrence) option ($0.30) and an adjusting immediate option that is certain (initially $0.15). All choice consequences are experienced during the assessment period, including monetary earnings delivered from a coin dispenser. Because the EDT involves real-time choice consequences, it is sensitive to state changes in discounting and is appropriate for use with children.
As a measure of temporal discounting, the EDT exhibits validity in several ways. Specifically, principal component analysis suggests the EDT loads on the same component as other recognized measures of discounting (Reynolds, et al., 2008), discounting trajectories are similar to those identified in other measures of CI (i.e., discounting curves from the EDT are more hyperbolic than exponential in shape (Reynolds & Schiffbauer, 2004)), and the EDT differentiates drug-using versus non-using groups (Reynolds, et al., 2006). Also, methylphenidate (a stimulant medication used for treatment of ADHD) decreased discounting on the EDT in children diagnosed with ADHD (Shiels, et al., 2009), whereas dopamine agonists increased discounting on the EDT (Barake, et al., 2013; Voon, et al., 2010). Additionally, administration of acute alcohol also increased delay discounting on the EDT (Reynolds, et al., 2006). However, it has also been argued that the EDT may link closely with other constructs (e.g., boredom proneness, probability matching) as opposed to delay discounting per se (Smits, et al., 2013).

Little research has evaluated test-retest reliability of discounting for the EDT. One small study of 26 participants suggests reliability is modest, with a test-retest correlation of .32 ($p = .11$) for assessments given over a seven-day period (Smits, et al., 2013). Further research is needed to more thoroughly evaluate reliability for the EDT.

**Monetary Choice Questionnaire**

The Monetary Choice Questionnaire (MCQ) (Kirby, et al., 1999) is a 27-item paper-and-pencil discounting assessment (an initial version of the MCQ had 21 items) (Kirby & Markovic, 1996). Each item requires choice between a larger-later sum of money available following a delay (7-186 days) and a smaller-sooner sum of money available immediately. The 27 items are
divided into 3 magnitude conditions: small ($25-35), medium ($50-60), and large ($75-85), allowing for calculation of a discounting parameter for each magnitude condition.

The MCQ was developed by assuming a hyperbolic model (Mazur, 1987), and determining the values of smaller-sooner, larger-later, and the delay to the larger-later based on specified discount rates (\(k\)) that range from a low of .00016 and a high of .25. The MCQ is scored by assigning a value equal to the geometric mean of the adjacent discount rates resulting in the highest proportion of consistent responses across each magnitude condition; thus, 10 discount rates (including the 2 endpoint values) are possible.

The MCQ has test-retest reliability of .71 across one year (Kirby, 2009) and is associated with impulsivity-related outcomes such as initiation of drug use (Audrain-McGovern, et al., 2009; MacKillop, et al., 2011), drug demand (Field, et al., 2006; MacKillop, et al., 2010) and craving (Giordano, et al., 2002). Another primary strength of the MCQ is the ease and brevity of administration. As discussed below, one limitation of the MCQ is that it may be less sensitive in the prediction of treatment outcomes than other measures of delay discounting (e.g., the EDT) (Krishnan-Sarin, et al., 2007; Reynolds, et al., 2006; Shiels, et al., 2009). Other limitations include the potential for ceiling and floor effects and limited modifiability while preserving task characteristics.

Task Considerations: Real vs. Hypothetical Rewards and Delays

Regardless of which human ITCT is used, whether tasks using hypothetical rewards and delays yield results comparable to those using real reward and delay procedures has been questioned. Real reward assessments may involve randomly selecting and honoring one choice
from all the choices the participant made during the CI assessment, thus resulting in either a delayed or immediate reward depending on which choice is selected by the experimenter. Some evidence indicates that discounting rates are steeper when using a real reward procedure compared to purely hypothetical rewards and delays (Kirby, 1997), although the reviewed studies had considerable differences in methods and outcome amounts used across studies (Lawyer, et al., 2011). Several studies directly comparing the two reward types reported no differences in discounting rates for real and hypothetical rewards (Baker, 2003; Johnson & Bickel, 2002; Lawyer, et al., 2011; Madden, et al., 2003; Madden, et al., 2004), with one study reporting that discounting rates involving real and hypothetical rewards are statistically equivalent (Matusiewicz, et al., 2013) and another reporting comparable outcomes at 4 different monetary amounts ($10, $25, $100, and $250) (Johnson & Bickel, 2002). CI assessed using hypothetical rewards was associated with real-world economic behavior, a finding that speaks to the predictive validity of hypothetical rewards (Bickel, et al., 2010). Furthermore, a neuroimaging study revealed that brain activation associated with discounting did not differ depending on whether real or hypothetical rewards were used in the task (Bickel, et al., 2009). Discounting for both types of rewards was associated with activation in an executive function area, the lateral prefrontal cortex, as well in reward-relevant limbic areas, including the anterior cingulate, posterior cingulate, and striatum. However, it is important to note that the realistic element in most of these “real reward” tasks was limited, with only one choice or a subset of choices paid to the participant, so probability may also been a factor influencing responding.

While many lines of research suggest that tasks using real and hypothetical rewards have a high level of concurrent validity, different relationships with the two types of tasks also have been reported, particularly with respect to effects of pharmacological manipulations and to
specific clinical populations (Acheson & de Wit, 2008; Acheson, et al., 2006; de Wit, 2009; McDonald, et al., 2003; Mitchell & Wilson, 2012; Paloyelis, et al., 2010; Reynolds, et al., 2006; Richards, et al., 1999). Acute alcohol increased choice impulsivity on the EDT, which differs from the “real reward” assessments described above in that rewards and delays are experienced by the participant for every choice as opposed to one randomly selected choice, but had no effect on the Richards delay-discounting task, which is more hypothetical in nature, as only a subset of choices are rewarded in the Richards task at the conclusion of the experiment of sometime thereafter (Reynolds, et al., 2006). Similarly, the EDT was sensitive to stimulant medication effects in children diagnosed with ADHD while the Richards task was not (Shiels, et al., 2009). However, when comparing participants with the combined subtype of ADHD (ADHD-CT) to healthy controls, the ADHD-CT group had a higher level of CI only when assessed by the hypothetical delay-discounting task, but not when assessed by a real-time delay-discounting task (Paloyelis, et al., 2010). Furthermore, variation in dopamine-related genes (i.e., COMT and DAT1) was related to discounting rates in hypothetical tasks, but not in real-time tasks (Paloyelis, et al., 2010). However, in a separate line of research, discounting on a delay-discounting task using hypothetical rewards was not related to two dopamine D2 receptor gene polymorphisms (i.e., ANKK1 Taq1A and C957T), although CI was increased under stress only in individuals with a specific dopamine D2 receptor polymorphism genotype (i.e., the CC genotype of the C957T polymorphism) (White, et al., 2009; White, et al., 2008). Amongst adolescent smokers, discounting on the EDT was not related to discounting on the MCQ, and scores on the former but not the latter related prospectively to treatment outcome measures (Krishnan-Sarin, et al., 2007).

While findings from research comparing real and hypothetical rewards have yielded divergent results, consideration of each of the studies separately is subject to differences in
sample size and characteristics. A quantitative meta-analytic examination of CI research would provide more conclusive evidence about the convergent validity of research using real and hypothetical rewards. In a large meta-analysis that included delay discounting research with real and hypothetical rewards, there was no evidence for differences in convergent validity between the two reward types (Duckworth & Kern, 2011). Given this finding, it is reasonable to conclude that the inclusion of both types of rewards in CI research may not be necessary, particularly when the assessment battery for a study is already large or burdensome.

**Health, Money, and Other Commodities**

The majority of human CI studies use measures that focus on choices between two financial outcomes (e.g., the MCQ). The popularity of this approach may be attributed to its robust reliability and validity across a range of clinical and non-clinical populations. An advantage is its concrete and quantifiable properties (therefore making these measures easily scalable) that make it both easily understandable across human populations and transferable to everyday behavior in modern society that relies upon money and financial decisions. Further, as a generalized conditioned reinforcer, monetary stimuli may be more useful for studies seeking a trait measure of CI that may be used to predict and explain an individual’s stable pattern of behavior across ranges of contexts and time. Indeed, research has shown predictive validity between financial measures of CI and behaviors ranging from substance use and addiction (Bickel, et al., 2013) to binge-eating disorder and obesity (Davis, et al., 2010; Fields, et al., 2013) to pathological gambling (Alessi & Petry, 2003). Finally, while robust outcomes have been reported in multiple Western cultures, its use in other cultures is less well reported and therefore should be approached cautiously. A few cross-cultural studies using monetary stimuli have reported both similar performance of the hyperbolic discounting equation as well as expected
cultural differences in the rate, with Western cultures typically discounting delayed rewards more steeply than Eastern cultures (Du, et al., 2002; Kim, et al., 2012) and in one study, showing greater activation of the ventral striatum when discounting future rewards (Kim, et al., 2012).

While money is the most commonly reported task stimulus, multiple CI tasks employing non-monetary commodities have shown good test-retest reliability and validity. The array of stimuli includes both primary natural reinforcers such as food (Estle, et al., 2007; Rasmussen, et al., 2010), sexual activity (Johnson & Bruner, 2013; Lawyer, et al., 2011; Lawyer & Schoepflin, 2013; Lawyer, et al., 2010) alcohol, and other substances of abuse (Bickel, et al., 2013; Odum, 2011), and conditioned, secondary reinforcers including music CDs, books and DVDs (Charlton & Fantino, 2008), availability of social interactions (Charlton, et al., 2013) and more abstract hypothetical concepts such as various types of health and environmental outcomes (Baker, 2003; Hardisty & Weber, 2009; Miller & Chapman, 2001).

In general, while different discounting rates have been obtained across different commodities (Charlton & Fantino, 2008; Odum, 2011), the degree of discounting between hypothetical commodities are highly correlated (Odum, 2011), and the same temporal discounting equations have been found to work well with these non-monetary commodities, with the hyperbolic function most commonly recommended (Charlton, et al., 2013). For more abstract outcomes such as delayed health and environmental outcomes associated with behavioral choices, framing appears to be particularly important, with delayed gains typically discounted at higher rates than delayed losses (Baker, 2003; Hardisty & Weber, 2009; Miller & Chapman, 2001; Mitchell & Wilson, 2010). Commodities that are able to be immediately consumed and are directly metabolized, such as food, alcohol and substances, tend to be temporally discounted at higher rates than commodities that serve more of an exchange function, such as music, books,
and money (Charlton & Fantino, 2008). This domain effect, or specificity effect, extends to observations that substance users discount their particular substance of abuse more so than monetary outcomes (Bickel, et al., 2013; Odum, 2011), and specific commodities are typically more closely related to a congruent domain-specific behavior than other commodities. For example, choice tasks using sexual activity and food stimuli were more predictive of sexual behavior and body fat percentage, respectively, than monetary stimuli (Lawyer & Schoepflin, 2013; Rasmussen, et al., 2010). As such, studies focused on particular clinical populations or behaviors of interest (e.g., cocaine use or binge-eating) may do well to supplement a generalized monetary trait measure of CI with a domain-specific task (e.g., cocaine or food stimuli, respectively) to maximize their ability to discriminate between populations and predict relevant behaviors, and to better disentangle potential roles of temperament and conditioned learning in those behaviors.

**Choice Impulsivity Research in Special Populations**

As described above, CI is considered an important aspect of psychiatric disorders (e.g., ADHD, addictions). CI also is a key feature of personality disorders (particularly cluster-B disorders like borderline and antisocial personality disorders). Additionally, CI patterns change across the lifespan (Christakou, et al., 2011). This section considers how best to implement and utilize ITCTs and self-reported measures of CI in clinical settings and with other special populations. This section reviews task features and participant characteristics that could influence results, the relative merits of task standardization versus modification and the potential clinical utility and treatment implications of CI research.
Task and Participant Considerations in Special Populations

*Hypothetical versus experiential.* As described previously, although results from hypothetical and experiential CI tasks correlate, this pattern may not be uniform across populations. For example, experiential tasks may be more suitable for those with difficulties with abstract reasoning and while including a probability component in a task may increase face validity, the increased cognitive load should be considered.

*Reward Modality.* As described previously, the reward modality and its salience in the population under examination should be considered. For example, cross-commodity studies in smokers suggest that drug choices are discounted more steeply than monetary rewards (Mitchell, 2004). Similarly, discounting rates appear to be somewhat commodity-dependent in cocaine users (i.e., $k$ is elevated for both money and cocaine, but higher for cocaine than for money) (Bickel, et al., 2011a). Differences between consumable (primary) reinforcers and non-consumable (secondary) reinforcers may also influence cross-commodity discounting in specific groups (Odum & Baumann, 2007).

*Rate dependence.* Baseline discounting rates, which may differ across clinical groups and ages (Bickel, et al., 2011a), are correlated with changes in delay discounting in clinical trials (Bickel, et al., 2007; Bickel & Yi, 2008). Given such differences, one might consider testing for regression to the mean or including the intercept or baseline discounting rate in models along with changes in delay discounting.

*Demographic factors: Age, socioeconomic status (SES) and intelligence.* CI is not uniform across the lifespan (Reimers, et al., 2009); delay discounting tends to be higher in childhood and adolescence (Casey & Jones, 2010; Odum, 2011), reduces as the prefrontal cortex
(PFC) develops in adulthood (Christakou, et al., 2011; Green, et al., 1996; Lockenhoff, et al., 2011; Samanez-Larkin, et al., 2011), and possibly increases again in older age (Harrison, et al., 2002) (although see Green et al., 1996; Lockenoff et al., 2011, Samanez-Larkin et al., 2011). In addition, SES and IQ are associated with multiple psychiatric disorders but are also independently associated with CI; thus, it may prove challenging to determine the individual contributions of SES and IQ to delay-discounting behaviors in certain clinical populations. Intelligence inversely correlates with delay discounting in both adults (de Wit, et al., 2007) and adolescents, independent of age (Olson, et al., 2007). In participants with IQs at or approaching the classification for an intellectual disability, this effect may in part result from limited understanding of the task. For this reason, a minimum IQ cut off should be used for study inclusion when possible. Experiential tasks may be more suitable for individuals with difficulties with abstraction, although as yet this has not been empirically tested. With regard to SES, an association of CI with socioeconomic status, income or education has been reported in some studies (de Wit, et al., 2007), although results from other studies have not supported this conclusion (Olson, et al., 2007). Subjects should be matched on these variables when groups are compared (Baker, 2003; Bickel, et al., 1999).

Temporal processes. It is thought that CI can be affected by time perception and temporal horizons (Ohmura, et al., 2005). Indeed, sleep-deprived participants underestimated intervals and had increased discounting compared to their own performances when in a non-sleep-deprived state (Reynolds & Schiffbauer, 2004). CI tasks may assume that temporal processes such as time perception and temporal horizons are uniform across populations. However, individual, cultural, and disorder-based differences may exist in envisioning the future (Kim, et al., 2012; Petry, et al., 1998; Teuscher & Mitchell, 2011). Although $k$ remained the most
important factor when time sensitivity was added to a two-parameter model of choice impulsivity (Jones, et al., 2009), such analyses have only been conducted in healthy (control) populations. It would therefore be pertinent to incorporate assessments of temporal processes in CI studies.

*Cigarette smoking.* While potential confounding effects of medications and illicit drugs on CI are commonly considered, cigarette smoking is often overlooked, despite data demonstrating increased discounting in smokers (Bickel, et al., 1999; Reynolds, 2006). This is particularly relevant when studying clinical groups who smoke at higher rates than the general population (Wing, et al., 2012).

*State versus trait effects.* CI is subject to state effects, such as sleep deprivation (Reynolds & Schiffbauer, 2004). Changes in mood also are associated with changes in CI (Weafer, et al., 2013). In addition, individual differences in delay-discounting responses to stress have been related to trait-like measures of perceived stress (Lempert, et al., 2012). Groups with high impulsivity may not be consistently impulsive, but rather may have moments of impulsive behavior in certain situations, disease states, while intoxicated, or while in drug withdrawal (Giordano, et al., 2002). Individuals with drug addictions made more impulsive decisions when presented with cues inducing craving (Dixon, et al., 2006). Certain groups (e.g., those with borderline personality disorder) may show greater CI under emotionally stressful circumstances. Thus, it is important to consider settings and context in CI research.

The above list is not meant to be exhaustive, and potential confounds should be considered for each research study, ideally at the design rather than analysis stage. Although removing all confounds may not be possible, they may be addressed by either experimental (i.e.,
matching) or statistical (i.e., covariation) control. A disadvantage of matching is that unrepresentative groups may be studied (e.g., patients with schizophrenia with above average IQs). Additionally, by systematically matching on a given measured variable one may be mismatching on an unmeasured variable (Meehl, 1971). However, despite these possibilities, one should also be cautious about covariation techniques (Miller & Chapman, 2001). Another option may involve using a relevant clinical control group. Qualitative research may also be used to reveal hidden confounders by providing insight into participants’ perspectives and unexpected task interpretation and strategies. Such approaches could be systematically incorporated into CI studies, but caution should be taken to avoid inadvertently influencing task results; alternatively, such efforts may be conducted in a pilot fashion.

**Standardization versus task modification**

Establishing the generalizability of CI measures is important for both between-group (e.g., comparing $k$ scores of individuals with and without substance-use disorders) and within-group studies (e.g., comparing $k$ across the lifespan). Several potential methods are available for testing the generalizability of CI results. In formal tests of measurement invariance, group or condition differences can be tested by comparing fit of constrained and unconstrained models. Tests of measurement invariance (Meredith, 2001) involve comparing models with increasingly constrained parameters (Widaman & Reise, 1997) and this approach is established within the psychometric literature to formally test comparability of assessments across groups (Borsboom, 2006). However, given measures of delay discounting may be derived from finding indifference points rather than estimating a latent variable from numerous manifest indicators, these tests of
invariance are essentially non-existent within the literature. Instead, some researchers have examined whether the form of discounting (e.g., a hyperbolic discounting function) is invariant across groups in order to bolster support that differences in discount rates are quantitative rather than qualitative. A study involving 935 individuals found the form of discounting is invariant (i.e., hyperbolic) across age, gender, ethnicity, IQ, or SES, suggesting differences in discounting in these groups may be attributable to quantitative rather than qualitative differences (Steinberg, et al., 2009). Additionally, researchers could utilize multiple types of CI tasks in order to estimate a latent variable representing a discounting rate to facilitate formal tests of measurement invariance.

If tasks are not equivalent across groups, researchers should consider the advantages (e.g., more appropriate for the test group and therefore perhaps more valid) and disadvantages (e.g., less comparable to healthy control group) of implementing a modified task within the context of the research question. For example, modifications could include a child-friendly interface and rewards, or utilizing an adapting paradigm, such as those used in neuroimaging studies, which base questions on subjects’ previous responses. While task equivalence is a prerequisite for any between-group comparisons, it may be advantageous to implement a modified task for a certain population in which group comparisons will be made (e.g., modifying a task for children, and then testing whether children with ADHD differ from children without ADHD). If possible, administering both standard and modified tasks would aid the accumulation of comparable data across studies, while still asking new and population-specific questions. Modifications should not be made lightly as the more data that are collected using standardized tasks, the more the results may be compared across studies.
Clinical utility and treatment implications

The association of CI with multiple clinical disorders (Reynolds, 2006) and clinically relevant phenomena like treatment outcome (Fernie, et al., 2013; Sheffer, et al., 2013; Stanger, et al., 2012) suggest that CI tasks could be applied in clinical settings. Potential uses include identification of patients at risk for future harmful behavior and as a treatment target and response indicator. Treatments such as contingency management may have potential in mitigating high impulsivity. Alternatively, cognitive (e.g., working-memory) training targeting the executive system has been shown to decrease delay discounting in cocaine users (Bickel, et al., 2011b) and improve alcohol and obesity treatment outcomes (Nederkoorn, et al., 2012; Verbeken, et al., 2013). However, it is important to note that psychiatric disorders like addictions are not synonymous with high impulsivity (Bechara, et al., 2001) and patients with ‘normal’ discounting rates may respond differentially to specific treatment strategies. Thus, CI tasks may help inform treatment personalization.

Despite the potential utility of CI tasks, research has largely been confined to group comparisons with less data on values that predict treatment outcome in individual patients. In order for CI tasks to develop into clinically useful tools, standard cut-off scores for dysfunctional $k$ or AUC values are important to establish. Although large-cohort CI studies have been conducted to meet this aim (Bickel, et al., 2012a; Reimers, et al., 2009), larger samples using standard tasks and taking into account variables such as age and education should be conducted in control, clinical and developmental groups (e.g., children, adolescents, older adults). In addition, as task scoring can be complicated, particularly for people in clinical settings without training, the development of a short, reliable and easy-to-administer task would facilitate clinical assessments as treatment settings are often characterized by time constraints.
Conclusions

CI is an important component of impulsivity in that it is: 1) reliably elevated in multiple, relevant patient populations; and, 2) translational, in that it can be measured in animals and humans.

CI is also unique; measures of CI show little if any correlation with other types of impulsivity measures, although there was moderate convergence between CI and trait impulsivity (as assessed by self-report and informant-report questionnaires) in a meta-analysis of impulsivity-related measures (Duckworth and Kern, 2011). With respect to rapid-response impulsivity, there are differences between populations regarding the two types of impulsivity, and the underlying neurobiology of rapid-response impulsivity and CI differs in preclinical studies. This supports the argument that CI measures should be included in any battery that is intended to assess the construct of impulsivity. With the body of research to date showing the importance of impulsivity in a variety of psychiatric disorders and the focus of the National Institute on Mental Health on Research Domain Criteria (RDoC) rather than specific psychiatric disorders (Insel, et al., 2010), having a comprehensive battery of measures to assess the construct of impulsivity is critically important. Such an approach is consonant with other efforts (e.g., PhenX) that seek to identify and characterize clinically relevant transdiagnostic measures.

There appears to be less evidence supporting which specific CI measures to include in a battery measuring impulsivity. Some investigators have argued that real rewards should be included in measures of impulsivity as opposed to hypothetical rewards. However, as discussed above, most data indicate that responding for real rewards is similar to responding for hypothetical rewards. Based on research that has shown that non-adjusting CI measures such as
the Monetary Choice Questionnaire and adjusting computer measures correlate but are not interchangeable (Epstein, et al., 2003), inclusion of one of each (adjusting and non-adjusting) would allow for the greatest comparison with prior studies. However, there are pros and cons to individual measures of CI that make the choice of specific measure less critical than the inclusion of a measure of CI in the battery. As discussed in this review, use of a standardized battery across studies would allow for more certainty regarding comparison of results across studies and subject populations. This was one of the goals of the PhenX Toolkit (https://www.phenxtoolkit.org) that now includes at least one measure of CI, the MCQ. One goal for the future of impulsivity research is a greater consensus on CI tasks that should be used in a standardized impulsivity battery. Important in this consideration may be the addition for specific conditions of disorder-specific commodities (food for obesity, drugs for addiction) in addition to monetary discounting. Additionally, the extent to which CI may reflect state-like or trait-like features warrants consideration and this may be examined through state manipulations (e.g., CI under stressed and non-stressed states). Other aspects of choice beyond magnitude and delay (involving risk, ambiguity, and other factors) may be manipulated experimentally and should be considered in ongoing and future studies (Tymula, et al., 2012). Considering the human conditions being modeled in animal studies is important, and efforts should be made to harmonize in as much as possible measures across species in order to facilitate cross-species comparisons and maximize translational impact of such research. With these points in mind, the field of research into CI and related constructs may be considered at a relatively early stage, and further refinements in recommendations are anticipated as the field matures further.

In summary, there was consensus among InSRI participants that at least one measure of CI should be included in any research battery of impulsivity. The details of which measure to
include and pros and cons of specific measures produced less consensus. However, all participants agreed that further research is warranted in this area, particularly in the area of the relationship between CI measures and clinical outcomes.

**Funding:** This research was supported in part by grants RL1 AA017539, R01 DA018647, R01AA021529, R01DA034755, R01DA030241, U19CA157345-0, R01-MH081181, R01 DA023087-01A2, R01 DA023476-01A2, P50DA09241 and P20DA027844 from the National Institutes of Health; the Connecticut State Department of Mental Health and Addiction Services; the Connecticut Mental Health Center; and a Center of Excellence in Gambling Research Award from the National Center for Responsible Gaming. The funding agencies did not provide input or comment on the content of the manuscript, and the content of the manuscript reflects the contributions and thoughts of the authors and do not necessarily reflect the views of the funding agencies.

**Conflict of Interest and Disclosures:** The authors report no conflicts of interest with respect to the content of this manuscript. Dr. Potenza has received financial support or compensation for the following: Dr. Potenza has consulted for and advised Somaxon, Boehringer Ingelheim, Lundbeck, Ironwood, Shire and INSYS; has received research support from the National Institutes of Health, Veteran’s Administration, Mohegan Sun Casino, the National Center for Responsible Gaming, and Forest Laboratories, Ortho-McNeil, Oy-Control/Biotie, Glaxo-SmithKline, and Psyadon pharmaceuticals; has participated in surveys, mailings or telephone consultations related to drug addiction, impulse control disorders or other health topics; has consulted for law offices and the federal public defender’s office in issues related to impulse control disorders; provides clinical care in the Connecticut Department of Mental Health and Addiction Services Problem Gambling Services Program; has performed grant reviews for the
National Institutes of Health and other agencies; has guest-edited journal sections and journals; has given academic lectures in grand rounds, CME events and other clinical or scientific venues; and has generated books or book chapters for publishers of mental health texts. Dr. Bickel is a principal of HealthSim LLC, has received support from the National Institutes of Health, and consulted with AstraZeneca and Prophase. Dr. Napier has received research support from the National Institutes of Health, the Michael J. Fox Foundation and the National Center for Responsible Gaming. Dr. Napier has received compensation for the following: consulting for a not-for-profit health education center and for law offices on issues related to addictions and impulse control disorders; speaking on addictions at community town hall meetings, public high schools, community-based not-for-profits, and professional meetings of drug courts; providing grant reviews for the National Institutes of Health and other agencies; and academic lectures and grand rounds. Dr. Napier is a member of the Illinois Alliance on Problem Gambling, and she provides expert advice on medication development to the Cures Within Research Foundation. Ms. Tedford has received support for training and research from the National Institutes of Health and from Rush University Graduate School. Dr. Mathias has no conflicts of interest to declare. Dr. Moeller is a consultant for Boehringer Ingelheim.
REFERENCES


of delayed rewards in attention deficit/hyperactivity disorder. Exp Clin Psychopharmacol, 17, 291-301.


<table>
<thead>
<tr>
<th>Human CI Task</th>
<th>Adjusting Amount</th>
<th>Internal Validity</th>
<th>External Validity</th>
<th>Construct Validity</th>
<th>Discriminant Validity</th>
<th>Reliability</th>
<th>Translational non-human analog</th>
<th>Overall Strengths</th>
<th>Overall Weaknesses</th>
<th>InSRI recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Richards</td>
<td>YES</td>
<td>Strong</td>
<td>Strong</td>
<td>Strong</td>
<td>Strong</td>
<td>Strong</td>
<td>YES</td>
<td>Widely used; Est links to neurobiology and clinical outcomes; OK for repeated measures</td>
<td>Possible insensitivity to experimental manipulations</td>
<td>YES</td>
</tr>
<tr>
<td>Monetary Choice Questionnaire</td>
<td>NO</td>
<td>Strong</td>
<td>Strong</td>
<td>Modest</td>
<td>Strong</td>
<td>Strong</td>
<td>YES</td>
<td>Widely used; Est links to neurobiology and clinical outcomes; OK for repeated measures; Sensitivity to manipulations; ease and brevity of administration</td>
<td>Less sensitive in predicting treatment outcomes; potential for ceiling and floor effects, limited modifiability</td>
<td>YES</td>
</tr>
<tr>
<td>Experiential Discounting Task</td>
<td>NO</td>
<td>Strong</td>
<td>Strong</td>
<td>Modest</td>
<td>Strong</td>
<td>Poor</td>
<td>NO</td>
<td>Widely used; Est links to clinical outcomes; OK for repeated measures; Sensitive to manipulations; variations used in children</td>
<td>Unable to use in fMRI scanner</td>
<td>YES</td>
</tr>
</tbody>
</table>