Published in final edited form as:
Soc Behav Pers. 2013 ; 41(3): 477–486. doi:10.2224/sbp.2013.41.3.477.

Assessment of risk-taking and impulsive behaviors: A comparison between three tasks

Sihua Xu[1,2], **Marc Korczykowski**[3], **Senhua Zhu**[1,3], and **Hengyi Rao**[1,3,*]

[1] Department of Psychology, Sun Yat-sen University, Guangzhou, China

[2] Department of applied sociology and psychology, Guangdong University of Business Studies, Guangzhou, China

[3] Center for Functional Neuroimaging, Department of Neurology, University of Pennsylvania, Philadelphia, PA, USA

Abstract

The balloon analogue risk task (BART), the delay discounting task (DDT), and the Iowa gambling task (IGT) are increasingly used for the assessment of risk-taking and impulsive behaviors. This study examined the reliability of and relationships between these three tasks in healthy Chinese subjects. The BART and DDT showed moderate to high test-retest reliability across three test sessions. However, the IGT showed low reliability for the first two sessions but high reliability for the last two sessions. Between tasks, only the BART and IGT showed significant correlations at the last two sessions, while no other correlations were found. These findings support the view that impulsivity is a complex construct with no single personality trait underlying the disposition for impulsive behaviors.

Keywords

Risk-taking; impulsivity; decision-making; Balloon Analogue Risk Task (BART); Iowa Gambling Task (IGT); Delay Discounting Task (DDT)

Impulsivity is a complex construct related to risk taking and lack of reflection between environmental stimuli and individual responding during decision making (Doob, 1990). Excessive risk-taking and impulsive behaviors are linked to many neuropsychiatric disorders such as drug addiction (Bechara, 2003; de Wit, 2009). Multiple self-report psychometric instruments have been developed and widely used to assess impulsive personality traits and behaviors, including the Eysenck Impulsiveness Scale (Eysenck, Eysenck, & Barrett, 1985) and the Sensation Seeking Scale (Zuckerman, 1994). However, the validity of self-report questionnaires relies heavily on the reading level and veracity of subjects.

Several cognitive tasks have also been developed and increasingly used to assess risk-taking and impulsive behaviors. These cognitive paradigms, including the balloon analogue risk task (BART), the delay discounting task (DDT), and the Iowa gambling task (IGT), focused more on the specific cognitive processes after controlling for situational factors (Cyders & Coskunpinar, 2011). For example, while initially designed to test patients with prefrontal lesions, the IGT has been widely used to measure risk-taking or impulsive behaviors in both healthy subjects and clinical populations (e.g., Aklin, Lejuez, Zvolensky, Kahler, & Gwadz,

[*] Correspondence should be addressed to: Hengyi Rao, Ph.D. Center for Functional Neuroimaging, Department of Neurology University of Pennsylvania 3400 Spruce Street, Philadelphia, PA 19104 USA Phone: (01)215-662-7341 Fax: (01)215-349-8260 hengyi@mail.med.upenn.edu.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

2005; Bechara, 2003; Lawrence et al., 2006). However, there remains a paucity of data regarding the reliability of the IGT, even in the clinical manual (Bechara, 2007). To our knowledge, there was only one published study which repeatedly used the IGT on two separate occasions (Verdejo-Garcia et al., 2007). However, this study reported significant learning effects in the IGT with no data on test-retest correlations. This knowledge gap may decrease further utility of the IGT for measuring impulsive behavior over time, for example, in patients before and after treatment interventions.

The BART is a cognitive paradigm for measuring behavioral risk-taking (Lejuez et al., 2002; 2003). In this task, subjects sequentially inflate a virtual balloon that may grow larger or explode. Behavioral performance on the BART has been shown to correlate with real world risk-taking and impulsive behaviors including alcohol use, cigarette smoking, drug use, gambling, theft, and unsafe sex (Hunt, Hopko, Bare, Lejuez, & Robinson, 2005; Lejuez et al., 2003; Lejuez, Aklin, Zvolensky, & Pedulla, 2003; Lejuez et al., 2002). A recent study has reported high test-retest reliability for the BART ($r = .77$, White, Lejuez, & de Wit, 2008).

Compared to the IGT and BART, the test-retest reliability of the DDT has been well documented. For example, Simpson and Vuchinich (2000) report a very high test-retest reliability for a hypothetical monetary DDT ($r = .91$). Baker, Johnson, and Bickel (2003) showed high DDT test-retest reliability in both smokers and nonsmokers ($r = .71$~.90). More recently, Beck and Triplett (2009) also observed high test-retest correlations ($r = .64$~.70) for a paper and pencil temporal discounting task. However, most of these studies employed western subjects; only one study (Ohmura, Takahashi, Kitamura, & Wehr, 2006) reported moderate to high test-retest correlations ($r = .60$) for the DDT in a Japanese population.

Although the BART, DDT, and IGT have been increasingly used for the assessment of risk-taking and impulsive behaviors throughout the world, there remains a paucity of data regarding the comparisons of the behaviors measured by these tasks, especially in Asian populations. The aim of the present study is therefore to fill this knowledge gap. The present study tested the three tasks in a cohort of 40 Chinese healthy young subjects on three separate occasions, with approximately 2-week intervals between two successive testing sessions. This experimental design also provides an opportunity to examine potential learning effects across multiple testing sessions.

Hypotheses

Based on the studies reviewed here, we proposed the following three hypotheses. H1: The DDT and BART will show high test-retest reliability across three test sessions, which is comparable to the literature. H2: The IGT task will show significant learning effect, with high test-retest reliability only in the last two test sessions after sufficient practice; H3: There will be significant correlations between the BART and IGT, but not between the BART and DDT.

Materials and Methods

Participants

A total of 40 healthy undergraduate and graduate students (15 male, age range 19~22 years) participated in the study. All participants completed a psychiatric symptom checklist (SCL-90; Derogatis, 1983) and reported no history of neurological or psychiatric disorders or head trauma. Participants were also excluded if they reported that they had drunk alcohol or smoked tobacco during the last two weeks. This study was approved by the Sun Yat-Sen University ethics committee. Written informed consent was obtained from all participants in accord with the Declaration of Helsinki. Each participant was compensated the same amount for participating in this study.

Procedure

Participants were tested three times with 2-week intervals. During each test session, participants completed risk-taking and impulsive decision making tasks, including the BART, the DDT, and the IGT, as well as the Sensation Seeking Scale (SSS, Zuckerman, 1994) and the impulsivity subscale of Eysenck Impulsiveness Scale-V (EIS, Eysenck, Pearson, Easting, & Allsopp, 1985). The order of the three tasks in each session was counterbalanced across subjects.

Instruments

The BART—During the BART, participants were required to press a button to sequentially inflate a series of 30 balloons which were displayed on the computer screen. The balloon could either grow larger or explode. A larger balloon was associated with an increased probability of explosion as well as a larger virtual monetary reward. Participants were instructed to maximize the amount of virtual reward during the experiment. Participants could discontinue balloon inflations at any time and win the reward for this balloon. If the balloon exploded, participants lost the reward. Risk-taking behavior on the BART was measured by calculating the mean adjusted number of balloon inflations, specifically, the average number of inflations on the trials when the balloon did not explode (Lejuez et al., 2003; Lejuez, Aklin, Zvolensky et al., 2003; Lejuez et al., 2002).

The DDT—During the DDT, participants were asked to select between two types of cards with virtual monetary reward: the left card with an immediate but small reward and the right card with a delay but large reward. Two possible delay rewards (large: ¥10,000, small: ¥1,000) and eight possible time intervals (6 hours, 1 day, 1 week, 2 months, 6 months, 1 year, 5 years and 25 years) were used for this task. Decision making behavior on the DDT was measured by the delay discounting rate (k), which was calculated by fitting the data to a hyperbolic decay model. These k values were Log transformed before statistical analysis.

The IGT—During the IGT, participants were required to choose one card from four decks of cards. Two decks were advantageous with overall a net win but smaller gain, while the other two decks were disadvantageous with larger gain but overall a net loss. The subjects need to complete 100 cards with no more than 60 cards for each deck (Bechara et al., 2001). Decks categorized as advantageous and disadvantageous were randomized for each test. Decision making behavior on the IGT was measured by the net scores of participants' choices, specifically, the number of selections from advantageous decks minus the number of selections from disadvantageous decks.

Impulsivity Scales

Participants also completed the EIS and SSS scales. The EIS has an internal consistency of .84 (Eysenck, Pearson et al., 1985) and the SSS has an internal consistency ranging from .83 to.86 (Zuckerman, 1994).

Data analysis

The stability of risk-taking and impulsive behaviors were assessed using one-way ANOVA, Pearson correlations, Spearman rank-order correlations, and intra-class correlation (ICC) analyses. Pearson correlations were also applied to examine the relationship between these tasks for each test session.

Results

Task Performance

Table 1 lists the mean risk-taking and impulsive behavioral performance from each task and questionnaires. Repeated ANOVA analyses revealed no differences between the three test sessions (all *ps* > .05).

Within-task Correlations

Table 2 lists the correlation coefficients for each of the three tasks across three sessions. Moderate to high correlations were observed for the BART and DDT across three test sessions. However, correlations for the IGT were only fair across the first two test sessions. Correlation coefficients did not differ between the DDT and the BART (all *ps* > .10), however, the Pearson and ICC correlation coefficients for the IGT were significantly lower than the BART (both *ps* < .05), and the Spearman rank-order correlation coefficient for the IGT was significantly lower than the large reward DDT ($p < .05$). In contrast, moderate to high correlations were observed for the IGT and no differences were found between the three tasks for the last two test sessions. There correlations significantly increased for the IGT from the first two test sessions to the last two sessions ($p < .05$).

Between-task Correlations

Table 3 lists the between-task correlation coefficients. For the first test session, there were no correlations between the three tasks. However, for the second and third test sessions, significant correlations were observed between the IGT and BART (both *ps* < .05), while the DDT showed no correlations with the IGT or BART. Significant, positive correlations were also found between the EIS and SSS scores across the three test sessions. However, no correlations were observed between task performances and psychometric scores (all *ps* > .05).

Discussion

There were several major findings from this study. First, we found that behavioral performance on the BART and DDT did not differ and both tasks showed moderate to high correlations across three separate test sessions. These findings are comparable to and consistent with the literature (Baker et al., 2003; Beck & Triplett, 2009; Simpson & Vuchinich, 2000; White et al., 2008), supporting the utility and feasibility of using these two tasks for the assessment of impulsive behaviors in Chinese populations.

Second, reliability of the IGT was significantly lower than those observed for the BART and DDT for the first two tests. These findings suggest that individual differences in impulsive behaviors measured by the BART and DDT are more stable and reproducible than the IGT. This finding is informative and important with respect to extending the utility of the IGT as a clinical instrument for assessing impulsivity and decision making deficits (Bechara, 2007). The findings also indicate that the stability of impulsive behavior measured by a single IGT test using 100 cards may not be acceptable in the absence of sufficient practice.

However, unlike a previous study (Verdejo-Garcia et al., 2007) showing significant IGT score increases, we found no significant changes in IGT behavior across three test sessions. This inconsistency may be explained by several factors, including the young subject cohort in the present study. It has been suggested that impulsivity and sensation seeking both decrease with age as individuals advance from adolescence to adulthood (Steinberg et al., 2008). Moreover, the monetary rewards were hypothetical and the subjects' actual participation compensation was fixed in our study. Therefore, subjects could have made more risky and impulsive decisions and chose more disadvantageous cards (Hinvest & Anderson, 2010). The observed overall negative mean net scores support this hypothesis. Finally, the Verdejo-Garcia et al. (2007) study employed a short time interval of 25 minutes. However, the present study used time intervals of two weeks between sessions, potentially increasing the ecological validity of the findings. Nevertheless, the non-significant trend of net score increases across three test sessions in this study are consistent with the between-session learning effect observed during the course of the IGT.

Third, we failed to find correlations between the IGT and BART at the first test session. However, we did observe significant correlations for last two test sessions, which is consistent with a recent study showing that performance on these two tasks correlated only in the later stages of the IGT (Upton, Bishara, Ahn, & Stout, 2011). The persistent correlations observed during the last two test sessions may reflect construct validity in impulsivity between the IGT and BART after controlling for learning effects. Future studies are needed to further examine the discriminative validity of risk-taking and impulsive behaviors measured by the IGT and BART.

In contrast, no correlations were found between the DDT and the BART or IGT. Similar null correlations have been reported (Dom, de Wilde, Hulstijn, & Sabbe, 2007; Reynolds, Ortengren, Richards, & de Wit, 2006), suggesting that risk-related cognitive processing mediating the BART or IGT behavior may be different than the DDT. Delay discounting is an assessment of the degree to which the subjective value of a commodity decreases as a function of a delay in its delivery (Rachlin, 2000). However, the BART conceptualizes risk taking propensity on a continuum in which risk is problematic only after a certain point, and that point varies across different balloon trials. In the IGT, the advantageous card decks may appear less attractive at first due to smaller wins, and subjects need to learn throughout the course of the task to discover the cards with overall positive net earnings. The net scores of the IGT therefore may reflect a combination of cognitive processes including learning speed and risk taking preference.

Finally, we failed to find any correlations between the tasks and the impulsivity scores. Several previous studies on different populations also reported no correlations between self-report impulsivity measures and behavioral performance from decision making tasks (Crean, de Wit, & Richards, 2000; Monterosso, Ehrman, Napier, O'Brien, & Childress, 2001; Reynolds, 2006). These findings suggest that behavioral tasks and self-report measures may tap different aspects of impulsivity.

There are several limitations in the present study. First, we used the hypothetical monetary reward and compensation was independent of subjects' task performance. The subjects may not be fully motivated to maximize their performance. Second, we employed a relatively small sample size of 40 young students with a narrow age range of 19~22 years. The prefrontal cortex may not be fully matured and greater risk taking and impulsivity have been reported in individuals of this age (Romer, 2010; Steinberg et al., 2008). Therefore, it remains unclear whether older adult will show a similar pattern of results. Further studies with a larger and broader cross-sectional sample of adults are necessary to replicate and extend the present findings. Finally, we only used fixed 2-week test-retest intervals. Longer time intervals may be valuable with respect to establishing the long-term stability and reliability of these tasks.

In summary, the present study provides important preliminary data in young Chinese adult samples regarding the reliability of these widely used impulsivity paradigms. Our results demonstrate that the BART and DDT tasks have moderate to high reliability for the assessment of risk-taking and impulsive behaviors. However, the reliability of the IGT may not be acceptable in the absence of sufficient practice. In addition, we only found significant correlations between the BART and the IGT in the second and third test sessions but not at the first test session, while no correlations were found between task performance and impulsivity scales. These findings support the view that impulsivity is a complex construct with no single personality trait underlying the disposition for impulsive behaviors.

Acknowledgments

This research was supported in part by Chinese NSF Grants 31070984 and 91124028, and NIH Grants R03 DA027098 and R21 DA032022.

References

Aklin WM, Lejuez CW, Zvolensky MJ, Kahler CW, Gwadz M. Evaluation of behavioral measures of risk taking propensity with inner city adolescents. Behaviour Research and Therapy. 2005; 43:215–228. doi: .1016/j.brat.2003.12.007. [PubMed: 15629751]

Baker F, Johnson MW, Bickel WK. Delay discounting in current and never-before cigarette smokers: similarities and differences across commodity, sign, and magnitude. Journal of Abnormal Psychology. 2003; 112:382–392. doi: 0021-843X.112.3.382. [PubMed: 12943017]

Bechara A. Risky business: emotion, decision-making, and addiction. Journal of Gambling Studies. 2003; 19:23–51. doi: 10.1023/A:1021223113233. [PubMed: 12635539]

Bechara, A. Iowa gambling task professional manual. Psychological Assessment Resources; Lutz: 2007.

Bechara A, Dolan S, Denburg N, Hindes A, Anderson SW, Nathan PE. Decision-making deficits, linked to a dysfunctional ventromedial prefrontal cortex, revealed in alcohol and stimulant abusers. Neuropsychologia. 2001; 39:376–389. doi: 10.1016/S0028-3932(00)00136-6. [PubMed: 11164876]

Beck RC, Triplett MF. Test-retest reliability of a group-administered paper-pencil measure of delay discounting. Experimental and Clinical Psychopharmacology. 2009; 17:345–355. doi: 10.1037/a0017078. [PubMed: 19803634]

Crean JP, de Wit H, Richards JB. Reward discounting as a measure of impulsive behavior in a psychiatric outpatient population. Experimental and Clinical Psychopharmacology. 2000; 8:155–162. doi: 10.1037//1064-1297.8.2.155. [PubMed: 10843297]

Cyders MA, Coskunpinar A. Measurement of constructs using self-report and behavioral lab tasks: Is there overlap in nomothetic span and construct representation for impulsivity? Clinical Psychology Review. 2011; 31:965–982. doi: 10.1016/j.cpr.2011.06.001. [PubMed: 21733491]

Derogatis, L. SCL-90-R manual-II. Clinical Psychometric Research; Towson, MD: 1983.

de Wit H. Impulsivity as a determinant and consequence of drug use: A review of underlying processes. Addiction Biology. 2009; 14:22–31. doi: 10.1111/j.1369-1600.2008.00129.x. [PubMed: 18855805]

Dom G, de Wilde B, Hulstijn W, Sabbe B. Dimensions of impulsive behaviour in abstinent alcoholics. Personality and Indivdual Difference. 2007; 42:465–476. doi: 10.1016/j.paid.2006.08.007.

Doob, L. Hesitation: impulsivity and reflection. Greenwood Press; New York: 1990.

Eysenck SBG, Eysenck HJ, Barrett P. A revised version of the psychoticism scale. Personality and Indivdual Difference. 1985; 6:21–29. doi: 10.1016/01918869(85)90026-1.

Eysenck SBG, Pearson PR, Easting G, Allsopp JF. Age norms for impulsiveness, venturesomeness, and empathy in adults. Personality and Individual Differences. 1985; 6:613–619. doi: 10.1016/0191-8869(85)90011-X.

Hinvest NS, Anderson IM. The effects of real versus hypothetical reward on delay and probability discounting. Quarterly Journal of Experimental Psychology (Hove). 2010; 63:1072–1084. doi: 10.1080/17470210903276350.

Hunt MK, Hopko DR, Bare R, Lejuez CW, Robinson EV. Construct validity of the Balloon Analog Risk Task (BART): associations with psychopathy and impulsivity. Assessment. 2005; 12:416–428. doi: 10.1177/1073191105278740. [PubMed: 16244122]

Lawrence NS, Wooderson S, Mataix-Cols D, David R, Speckens A, Phillips ML. Decision making and set shifting impairments are associated with distinct symptom dimensions in obsessive-compulsive disorder. Neuropsychology. 2006; 20:409–419. doi: 10.1037/0894-4105.20.4.409. [PubMed: 16846259]

Lejuez CW, Aklin WM, Jones HA, Richards JB, Strong DR, Kahler CW, Read JP. The Balloon Analogue Risk Task (BART) differentiates smokers and nonsmokers. Experimental and Clinical Psychopharmacology. 2003; 11:26–33. doi: 10.1037/1064-1297.11.1.26. [PubMed: 12622341]

Lejuez CW, Aklin WM, Zvolensky MJ, Pedulla CM. Evaluation of the Balloon Analogue Risk Task (BART) as a predictor of adolescent real-world risk-taking behaviours. Journal of Adolescence. 2003; 26:475–479. doi:10.1016/S0140-1971(03)00036-8. [PubMed: 12887935]

Lejuez CW, Read JP, Kahler CW, Richards JB, Ramsey SE, Stuart GL, Brown RA. Evaluation of a behavioral measure of risk taking: the Balloon Analogue Risk Task (BART). Journal of Experimental Psychology: Applied. 2002; 8:75–84. doi: 10.1037/1076-898X.8.2.75. [PubMed: 12075692]

Monterosso J, Ehrman R, Napier KL, O'Brien CP, Childress AR. Three decision-making tasks in cocaine-dependent patients: do they measure the same construct? Addiction. 2001; 96:1825–1837. doi:10.1046/j.1360-0443.2001.9612182512.x. [PubMed: 11784475]

Ohmura Y, Takahashi T, Kitamura N, Wehr P. Three-month stability of delay and probability discounting measures. Experimental and Clinical Psychopharmacology. 2006; 14:318–328. doi: 10.1037/1064-1297.14.3.318. [PubMed: 16893275]

Rachlin, H. The Science of Self-control. Harvard University Press; Cambridge, MA: 2000.

Reynolds B. A review of delay-discounting research with humans: relations to drug use and gambling. Behavioural Pharmacology. 2006; 17:651–667. doi: 10.1097/FBP.0b013e3280115f99. [PubMed: 17110792]

Reynolds B, Ortengren A, Richards JB, de Wit H. Dimensions of impulsive behavior: Personality and behavioral measures. Personality and Individual Differences. 2006; 40:305–315. doi: 10.1037/1064-1297.16.2.124.

Romer D. Adolescent risk taking, impulsivity, and brain development: implications for prevention. Developmental Psychobiology. 2010; 52:263–276. doi: 10.1002/dev.20442. [PubMed: 20175097]

Simpson CA, Vuchinich RE. Reliability of a measure of temporal discounting. Psychological Record. 2000; 50:3–16.

Steinberg L, Albert D, Cauffman E, Banich M, Graham S, Woolard J. Age differences in sensation seeking and impulsivity as indexed by behavior and self-report: evidence for a dual systems model. Developmental Psychology. 2008; 44:1764–1778. doi: 10.1037/a0012955. [PubMed: 18999337]

Upton DJ, Bishara AJ, Ahn WY, Stout JC. Propensity for risk taking and trait impulsivity in the Iowa Gambling Task. Personality and Individual Difference. 2011; 50:492–495. doi: 10.1016/j.paid. 2010.11.013.

Verdejo-Garcia A, Benbrook A, Funderburk F, David P, Cadet JL, Bolla KI. The differential relationship between cocaine use and marijuana use on decision-making performance over repeat testing with the Iowa Gambling Task. Drug and Alcohol Dependence. 2007; 90:2–11. doi: 10.1016/j.drugalcdep.2007.02.004. [PubMed: 17367959]

White TL, Lejuez CW, de Wit H. Test-retest characteristics of the Balloon Analogue Risk Task (BART). Experimental and Clinical Psychopharmacology. 2008; 16:565–570. doi: 10.1037/a0014083. [PubMed: 19086777]

Zuckerman, M. Behavioral expressions and biosocial bases of sensation seeking. Cambridge University Press; New York: 1994.

Table 1

Mean decision making behavior of the three tasks as well as impulsivity scores at test sessions 1, 2 and 3. Standard deviations in parentheses.

	Session 1	Session 2	Session 3
BART	42.62 (14.50)	42.35 (14.04)	43.79 (14.21)
I G T	-11.63 (24.03)	-7.20 (25.58)	-7.90 (29.57)
L.DDT	-1.35 (1.06)	-1.60 (1.06)	-1.85 (1.24)
S. DDT	-1.36 (0.83)	-1.64 (1.15)	-1.68 (1.25)
SSS	18.03 (5.57)	18.83 (6.02)	18.43 (6.27)
EIS	7.00 (3.70)	7.05 (4.08)	7.33 (3.61)

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 2

Within-task correlations for the three tasks across test sessions 1, 2 and 3.

	Session 1 & 2			Session 2 & 3		
	Pearson-r	ICC-r	rho	Pearson-r	ICC-r	rho
BART	.66***	.66***	.60***	.76***	.76***	.73***
IGT	.35*	.35*	.36*	.65***	.65***	.51***
L.DDT	.55***	.53***	.68***	.70***	.67***	.73***
S.DDT	.55***	.51***	.63***	.80***	.80***	.76***

**p<.01

* p<.05

*** p<.001, two-tailed.

Table 3

Between-task Pearson correlations at three test sessions.

	BART	IGT	L.DDT	S.DDT	SSS	EIS
Session 1						
BART	1	-.25	.01	.07	.23	-.01
IGT		1	.28	.28	-.28	-.24
L.DDT			1	.77***	-.07	-.01
S.DDT				1	.08	-.05
SSS					1	.57**
EIS						1
Session 2						
BART	1	-.36*	-.12	-.11	.01	-.01
IGT		1	.22	.22	-.23	.19
L.DDT			1	.87***	.08	-.04
S.DDT				1	.15	.01
SSS					1	.75**
EIS						1
Session 3						
BART	1	-.34*	-.16	-.22	-.14	-.01
IGT		1	-.02	-.01	-.08	-.01
L.DDT			1	.91***	.07	-.12
S.DDT				1	.15	-.10
SSS					1	.67**
EIS						1

* $p<.05$

** $p<.01$

*** $p<.001$, two-tailed.