Published in final edited form as:

Exp Clin Psychopharmacol. 2014 October ; 22(5): 444–452. doi:10.1037/a0037421.

A New Look at Risk-Taking: Using a Translational Approach to Examine Risk-Taking Behavior on the Balloon Analogue Risk Task

Kelly S. DeMartini^{1,*}, Robert F. Leeman¹, William R. Corbin², Benjamin A. Toll^{1,3,4}, Lisa M. Fucito¹, Carl W. Lejuez⁵, and Stephanie S. O'Malley^{1,3}

¹Yale University School of Medicine, Department of Psychiatry, New Haven CT

²Arizona State University, Department of Psychology, Tempe AZ

³Yale Cancer Center, New Haven CT

⁴Smilow Cancer Hospital at Yale-New Haven, New Haven CT

⁵University of Maryland, Department of Psychology, College Park MD

Abstract

Models of risk-taking typically assume that the variability of outcomes is important in the likelihood of making a risky choice. In an animal model of the Balloon Analogue Risk Task (BART), within-session variability, or the coefficient of variability (CV), was found to be a novel predictor of behavior (Jentsch et al., 2010). Human studies have not investigated how BART performance differs when using the CV versus a traditional BART measure (e.g. number of pumps). This study sought to determine whether the CV provides a unique and valuable alternative index of risk-taking on the BART, and to determine the relationship of the CV to self-reported alcohol consumption. Young adult heavy drinkers ($n = 58$, 72% male, mean age 21.53) completed an assessment of drinking patterns and a modified version of the BART. Multiple regression results indicated that CV is a unique predictor of total explosions and total money earned on the BART. Higher levels of variability were associated with fewer explosions but less money earned, whereas more pumps was associated with more explosions but more money. Higher CV was also associated with lower lifetime and past 3 months peak drinking quantity, higher levels of self-efficacy to control drinking, and lower levels of drinking acceptability (i.e., injunctive norms). Total pumps was associated with higher lifetime peak drinking, lower self-efficacy to control drinking, and higher levels drinking acceptability. Overall, the CV can provide an alternative method of assessing BART performance and the association of risk-taking with drinking patterns.

*Corresponding Author: Kelly S. DeMartini, Ph.D., Yale University School of Medicine, Department of Psychiatry, Division of Substance Abuse, 1 Long Wharf Drive, Box 18, New Haven, CT 06511, (p) 203-974-5784, (f) 203-974-5790, kelly.demartini@yale.edu.

Authors Contribution

KSD, RFL, and SSO were responsible for the study design and concept. KSD performed the statistical analysis and drafted the manuscript. All authors assisted with interpretation of the findings, provided critical revision of the manuscript, and approve the final version for publication.

Keywords

alcohol; Balloon Analogue Risk Task (BART); coefficient of variability; risk taking; young adults

Risk-taking pervades nearly all aspects of human life. Any implementation of a goal-directed behavior can be considered a risk-taking behavior when the behavior can lead to more than one outcome and when some of those outcomes are undesirable and/or dangerous (Furby & Beyth-Marion, 1992). Most definitions of risk-taking behavior, therefore, concentrate on the opportunity to gain reward with a corresponding potential for loss or harm (e.g. Leigh, 1999). Because of the ubiquity of risk-taking behavior, numerous academic fields, including psychology, etymology, and economics have investigated theoretical models of risk (Weber, Shafir, & Blais, 2004).

Both human and animal studies of risk-taking behavior have proposed models that predict risky choice in either a deterministic model or predict risk sensitivity (i.e. the probability of choosing a risky or less risky option) in a stochastic model. Importantly, nearly all of these models assume that the variability of each option's outcomes is an important factor in the likelihood of choosing a risky option (Weber et al., 2004). As a result, the variance of outcomes is commonly used as the measure of the variability of outcomes. Yet, both human and animal behavior often deviate from model predictions (see Shafir et al., 1999). For example, both animals and people may perceive outcome variability as relative to the average level of outcomes, rather than in an absolute determination. These characteristics of subjective perception, which are unaccounted for in models that utilize outcome variance to measure risk, need to be considered to arrive at accurate behavioral predictions (Weber et al., 2004; Weber & Hsee, 1999). These findings, and others, have led to the suggestion that variance may not be the optimal measure of risk (Weber et al., 2004). Instead, the coefficient of variability (CV), a measure of the variability of risky choice alternatives calculated by dividing the standard deviation (SD) of outcomes by their expected value (EV; e.g. mean value), may predict risk sensitivity better than variance or the SD (Weber et al., 2004; Shafir, 2000). It also offers the unique advantage of allowing comparisons across analyses, because dividing the SD by the EV creates a dimensionless unit of measure.

The CV, however, is a relatively new index of risk. In a meta-analysis of animal models of foraging behavior, CV improved predictions of risk preference (Shafir, 2000). A meta-analysis of human risky choice studies showed that CV was a significant predictor of risky choice (Weber et al., 2004). Specifically, in gain domains, larger CV was associated with a greater proportion of participants choosing sure outcomes instead of risk outcomes. In loss domains, larger CV was associated with a lower proportion of choices of sure outcomes. This was particularly true when information about variability is acquired by experience, as in animal behavior (Weber et al., 2004). More recently, intra-individual variation in responding (e.g. CV) has been thought of as an index of the level of executive control an individual is able to employ. Using both a behavioral task of inhibition (e.g. the Go/No-go paradigm) and fMRI data, normal participants with higher intra-individual variability activated inhibitory brain regions, including the right inferior parietal and thalamic regions and bilateral middle frontal areas, to a greater extent than those who showed less variability.

Higher CV was correlated with poorer inhibitory performance (Bellgrove, Hester, & Garavan, 2004). That more neural activation is required may indicate that inhibition is more difficult for these individuals. Thus, using the CV as an indicator of risk may provide new information about risk-taking behavior and impulsivity.

That the CV can provide a new way of assessing performance on behavioral risk-taking tasks has recently been illustrated with a rodent version (Jentsch et al., 2010) of the Balloon Analogue Risk Task (BART; Lejuez et al., 2002). In the original, computerized BART, humans make decisions about the amount of risk they are willing to accept to obtain a monetary reward on each trial of the task. Participants sequentially inflate, with the click of a button, a balloon that could either explode or grow larger. They click one button successively to increase the earned reward, but each click has an associated risk of trial failure (e.g., a balloon may burst). On each trial, participants can continue accepting additional risk by continuing to click (e.g. increase the size of the balloon) or end the trial by accepting the money earned to that point. BART risk preferences are correlated with risk-related constructs, including sensation-seeking and impulsivity (Cf. Bornovalova et al., 2009; Lejuez et al., 2002), and risky behaviors, including smoking and alcohol use (Ashenhurst et al., 2011; Fernie et al., 2010; Larsen et al., 2013). In an animal model of the BART, rats continually pressed a lever for food or pressed a “cash out” lever. Continually pressing had a risk of trial failure and correspondent loss of food due to over-responding. The CV was used as a novel performance metric in a recent study using the animal model of the BART (Jentsch et al., 2010). Under no-risk conditions, CV was not a significant predictor of pellets earned. Under high-risk conditions, however, CV was negatively associated with pellets earned, and total lever presses did not explain food pellets earned. Rats with higher CV constrained their responding as risk increased and so earned fewer pellets. The authors concluded that high intra-subject, intra-session variability could indicate poor top-down behavioral control, which could have led to the higher CT rats’ inability to maximize reward receipt (Jentsch et al., 2010).

To date, no human study has investigated how patterns of responding on the BART differ when using the CV versus a traditional BART outcome measure (e.g. pumps; Lejuez et al., 2002). We sought to determine whether CV and pumps were differentially related to two measures of performance on the BART, explosions and money earned, in a clinical sample of young adult heavy drinkers. Based on pre-clinical findings (Jentsch et al., 2010), we hypothesized that CV would be negatively related to total money earned. The animal BART model did not include a measure of failed/incorrect trials that would compare to the human BART’s explosions variable. Therefore, based on the assumption that higher levels of variability would lead to more trials with higher levels of risk, we hypothesized that CV would be positively related to explosions. Similarly, based on pre-clinical data indicating that CV was negatively associated with food earned in high-risk situations (Jentsch et al., 2010), we were interested in behavior on trials immediately following exploded balloons. Because these trials might be construed as having more risk because of their chronological association with a failure, we also hypothesized that CV would be negatively associated with pumps on trials after exploded balloons. Lastly, based on past research with the BART, we hypothesized that pumps would be positively related to money earned and explosions. We also sought to determine whether CV and pumps differentially related to measures of

alcohol use and to known predictors of alcohol use; perceived self-efficacy to control drinking; protective behavioral strategy use; and descriptive and injunctive drinking norms. Ability to control drinking is associated with decreased consumption (Sitharthan & Kavanagh, 1991) and better inhibitory control is a component of general self-control (Muraven, 2010). Because higher CV has been posited to reflect poorer inhibitory control (Jentsch et al., 2010), we hypothesized a negative relationship between CV and ability to control drinking and a positive relationship between CV and alcohol use. The use of protective behavioral strategies, defined as cognitive-behavioral strategies used by individuals before or while drinking alcohol to limit negative alcohol-related consequences (Martens et al., 2004) is associated with decreased alcohol consumption and/or consequences (DeMartini et al., 2013). The ability to use these strategies, however, requires planning and/or effortful cognition; because higher CV may be associated with less cognitive control, we hypothesized a negative relationship between CV and protective behavioral strategies. Lastly, descriptive and injunctive norms are particularly important predictors of alcohol consumption in college students (see Baer, Stacy, & Larimer, 1991; Larimer et al., 2004; Neighbors, Lee, Lewis, Fossos, & Larimer, 2007). Alcohol interventions are more effective when they present information to students about norms (Carey, Scott-Sheldon, Carey, & DeMartini, 2007) and campus-wide marketing interventions have targeted drinking norms for this population (see Wechsler, Nelson, Lee, Seibring, Lewis, & Keeling, 2003). Despite being one of most studied factors to impact alcohol consumption, research has not directly investigated whether either type of norm is associated with risk-taking. Given that both types of norms, as well as impulsivity (Dick et al., 2010), are associated with increased alcohol use, we hypothesized a positive relationship between CV and descriptive and injunctive norms.

Method

Participants

Participants in the current study were recruited as potential participants in a larger clinical trial for young adult heavy drinkers. Young adults between ages 18 and 25 years were recruited via flyers, newspaper, television, and online advertisements (e.g. Facebook, Craigslist) to participate in an ongoing 8-week randomized, double-blind placebo-controlled clinical trial of naltrexone plus brief individual counseling to reduce the frequency of any alcohol consumption and frequency of heavy drinking. Compensation of up to \$500 was advertised and no explicit motivation to change drinking was required for trial inclusion. Inclusion criteria included heavy drinking (≥ 5 standard drinks for men and ≥ 4 standard drinks for women) on 4 or more days within the 28 days prior to intake appointment. Exclusion criteria included (1) a current DSM-IV diagnosis of alcohol dependence that was clinically severe (e.g., reported a Clinical Institute of Withdrawal Assessment scale [CIWA; Sullivan et al., 1989] score of ≥ 8), (2) a DSM-IV diagnosis of drug dependence other than nicotine dependence (including marijuana dependence) (3) a urine drug screen indicating the use of illicit substances, except marijuana, (4) serious psychiatric illness (e.g., schizophrenia, substantial suicide risk), and (5) current pregnancy or lactation.

Procedure

Prospective participants for the clinical trial of naltrexone completed a preliminary screener of eligibility by telephone or the internet. Those deemed likely to be eligible then attended an in-person intake appointment. At intake, all prospective participants provided written informed consent. Self-report assessments were completed on the internet at the intake appointment in our research location. Prospective participants also completed the BART (see below for description) at the research location. Forthcoming reports will describe treatment outcomes. This report focuses on self-report and risk-taking behavior gathered at the intake appointment from the main trial of the 58 young adults who had complete data on the BART. The trial was approved by the Institutional Review Board of the Yale University School of Medicine.

Measures

Risk-Taking Behavior

Risk-taking behavior was assessed with an adapted and computerized version of the Balloon Analogue Risk Task (BART; Lejuez et al., 2002). The computer screen showed a simulated balloon accompanied by a balloon pump, a reset button labeled *Collect \$\$\$*, a display indicating money earned permanently from previous trials labeled *Total Earned*, and a second display, labeled *Last Balloon*, that listed the money earned on the previous balloon.

In the original BART, participants click a button for each pump they choose to add to a balloon (Lejuez et al., 2002). In current version (e.g. the automatic BART), participants were asked to type the total number of pumps, out of a possible 128 for each balloon. This modified version allows participants to type in as large a number of pumps as they wish. Each pump was worth 1 cent, which participants were able to keep unless the number of pumps selected was greater than the explosion point. In this case, the balloon exploded, and no money was awarded for the balloon. After each explosion or money collection, a new balloon appeared. Participants completed 30 trials (e.g. balloons). Participants were not given information about the probability of an explosion, but were informed at the outset that inputting 64 pumps could provide optimal performance. In contrast to the original BART, the automatic BART, therefore, allows analysis of the number of pumps on all balloons, whereas the original only allows analysis of unexploded balloons (Cf. Lejuez et al., 2002). Therefore, the original version omits what could be some of the riskiest decisions (i.e., decisions that led directly to the negative outcome of balloon explosion). Research on the effect of these modifications, including providing participants with the expected-value-maximizing strategy, indicates that they produce unbiased BART statistics (Pleskac, Wallsten, Wang, & Lejuez, 2008). The original approach of analyzing only unexploded balloons could produce somewhat biased results. Specifically, scores were biased toward low scores, because the more times respondents choose a risky option, the more likely it is that a trial will end in failure (e.g. an explosion). Because of this, the adjusted score filters out longer response sequences and biases scores toward a lower number of pumps (Pleskac et al., 2008). Additionally, whereas participants were largely risk averse on the original BART, these modifications have been found to increase the number of risks that participants are willing to take. The automatic BART has also been found to retain similar relationships

to substance use indices as the standard version of the task (Pleskac et al., 2008). The standard BART and other, similar risky decision making tasks provide valuable information regarding within trial decision making (Pleskac 2008) that is absent from the automatic BART. Despite this disadvantage, overall, this updated version of the BART provides better estimates of risk-taking while maintaining its predictive utility.

Several outcome variables were calculated to examine participants' decision making. To examine overall risk-taking, mean *total pumps* was calculated by summing the total pumps per trial and dividing that number by 30 trials. While adjusted pumps is the main outcome measure with the standard version of the BART (Lejuez et al., 2002), given that it is possible to account for decisions made on all balloons in the automatic BART, total pumps becomes the standard outcome measure (Pleskac et al., 2008). *Total pumps by time-block* assessed changes in risk-taking over blocks of the BART (e.g. Lejuez et al., 2002). The 30 trials of the BART were divided into 3 blocks of 10 trials (i.e. the first 10 trials, the second 10 trials, and the last 10 trials) (Lejuez et al., 2002). Total pumps for each set of 10 trials were averaged together. Each person, therefore, had 4 scores on pumps: (1) mean total pumps over the 30 trials; (2) mean pumps for the first 10 trials; (3) mean pumps for the second 10 trials; and (4) mean pumps for the last 10 trials.

A similar procedure was used to assess BART performance. Each participant had 4 scores on explosions: (1) total explosions over the 30 trials; (2) total explosions during the first 10 trials; (3) total explosions during the second 10 trials; and (4) total explosions during the last 10 trials. Each participant also had 4 scores on money earned: (1) total money earned over the 30 trials; (2) total money during the first 10 trials; (3) total money during the second 10 trials; and (4) total money during the last 10 trials. Overall, explosions variables were used to examine lack of success on the BART and money-earned variables were used to examine success on the BART¹.

To determine whether variability in responding is related differently to BART performance than to traditional measures of risk-taking (e.g. *total pumps*), the total *coefficient of variability* was calculated for each participant with the following equation:

$$CV_{total} = SD[\text{total pumps across 30 trials}] / \text{Mean}[\text{total pumps across 30 trials}]$$

CV scores for each block of 10 trials were also calculated with the following equation:

$$CV_{time-block} = SD[\text{total pumps per 10 trials}] / \text{Mean}[\text{total pumps per 10 trials}]$$

To determine whether CV predicted how participants reacted after an exploded balloon, 3 scores were created: (1) mean pumps on exploded balloons; (2) mean pumps on balloons immediately after exploded balloons; and (3) the difference between exploded balloons and post-explosion balloons.

¹Traditional analyses of the BART often include analyses of total adjusted pumps, or pumps on unexploded balloons. In the version used in these analyses, money earned is equivalent to adjusted pumps. Money earned is the clicks used on unexploded balloons.

Alcohol Use

The Daily Drinking Questionnaire-Revised (DDQ-R; Collins, Parks, & Marlatt, 1985) assessed typical drinking in the 3 months before intake with two questions: 1) the number of times participants consumed any alcohol for each day of the week within the previous 13 weeks and 2) the number of standard drinks consumed on a typical drinking day. Drinks per drinking day (DDD) was calculated by taking the mean of the number of drinks typically consumed on each day of the week, weighted according to the number of days in the previous 13 weeks when drinking occurred. Peak drinking was assessed with 3 single item questions: 1) the largest number of drinks consumed on a single day within the past 3 months, 2) the frequency of drinking to the peak number of drinks using the following scale: 1 = 1–2 days in the past 3 months; 2 = 1 day a month; 3 = 2–3 times a month; 4 = 1 day a week; 5 = 2 days a week; 6 = 3–4 times a week; 7 = 5–6 times a week; 8 = every day, and 3) the largest number of drinks ever consumed on a single day in the participant's lifetime.

Alcohol-Related Constructs

Controlled Drinking Self-Efficacy was assessed with the Controlled Drinking Self-Efficacy Scale (CDSSES; Sitharthan et al., 2003). The CDSSES is a 20-item self-efficacy measure to assess confidence to reduce overall drinking consumption and frequency and drinking in response to social situations and negative affect. Confidence is rated on a 0 (0% confident) to 10 (100% scale). It has demonstrated high test-retest reliability ($r = 0.90$) and internal consistency (coefficient alpha = 0.95). *Protective Behavioral Strategies* was assessed with the Protective Strategies Questionnaire (PSQ; Palmer, 2004; DeMartini et al., 2013). The PSQ is a 10-item self-report measure to assess the frequency with which participants engage in protective strategies. All items are scored on a 1–7 Likert scale (1 = never, 7 = always). The PSQ has two factors, Direct Strategies (e.g. “space drinks over time”) and Indirect Strategies (e.g. “have a designated driver”), which have been validated across clinical and general university samples (DeMartini et al., 2013). *Descriptive drinking norms* were assessed with the Drinking Norms Rating Form (DNRF; Baer et al., 1991). The DNRF asks participants to indicate how much an average member of their social group drinks on each day of a typical week. A total score was created by summing the number of drinks indicated for each of the seven days. *Injunctive norms*, or views about others' perceptions of drinking behavior, were rated on 7 items. Each item assessed the extent to which participants believed the average member of their social group would agree or disagree with each statement (e.g. “Drinking is all right, but I should never get ‘smashed’”). Items were rated on a 5-point Likert scale (1 = disagree; 5 = agree). A total injunctive norms score was calculated by averaging all items.

Analytic Plan—Because the CV has rarely been used to study risk-taking behavior in human research, exploratory bivariate correlations were used to assess overall relationships among CV, explosions, and money earned on the BART. Bivariate correlations were then calculated among total pumps, explosions, and money earned to assess relationships between the traditional performance measure (total pumps) and the new measure (the CV). A scatter-plot of CV and pumps was created to examine the relationship between total pumps and CV. To further examine the relationship between BART performance and CV, participants were defined as being high CV or low CV. High CV participants were those

with mean CV values 1SD above the mean CV value. Low CV participants had CV values 1SD below the mean CV value. Between-groups *t*-tests were used to compare the groups on pumps, explosions, and money earned. For illustrative purposes, Figure 1 displays a scatter-plot of pumps per trial for a high CV participant (at least 1 SD above the sample mean) and a low CV participant (at least 1 SD below the sample mean).

Multiple regression models were used to determine whether the CV was a unique predictor of money earned and explosions. To predict money earned and explosions in separate models, pumps and CV were simultaneously entered as independent variables. Multiple regression models were also used to examine whether CV was related to risk-taking behavior after a negative consequence (e.g. an exploded balloon). Regressions examined whether CV predicted how participants reacted after an exploded balloon. CV was entered as the sole predictor of the difference score of pumps (mean pumps on exploded balloons - mean pumps on balloons immediately after exploded balloons). Sex was included as a covariate in the models.

A series of three repeated measures ANOVAs was used to examine how risk-taking (e.g. pumps) and success (e.g. explosions and money earned) changed over blocks of the BART trials. Analysis by blocks of 10 trials is a standard way of examining changes in BART performance (see Lejuez et al., 2002 for an example). A repeated measures ANOVA was also used to examine whether variability in risk-taking (e.g. CV) changed over the BART trials. Partial eta squared (η^2_p) describes the total variability of the dependent variable that is attributable to the effect. Following significant within-subjects omnibus tests, trend analysis was completed with polynomial contrasts to determine the functional relationship between time (e.g. repeated BART trials) and performance on the BART (Keppel & Wickens, 2004). Linear and quadratic trends were tested. Lastly, bivariate correlations examined the relations between variability in responding on the BART and alcohol use, controlled drinking self-efficacy, and injunctive norms.

Results

Sample Characteristics

The overall sample had a mean age of 21.53 (SD = 2.11) years and was predominantly male ($n = 42$, 72%) and Caucasian ($n = 44$, 76%). Most participants had completed some college ($n = 25$, 43%) or had a college degree ($n = 14$, 24%). Participants reported drinking an average of 6.02 drinks per drinking day (SD = 3.29), an average of 16.41 (SD = 6.67) drinks on their peak occasion in the past 3 months, and an average of 20.47 (SD = 7.62) drinks on their peak lifetime occasion.

Relationships among CV, Pumps, and BART Performance

Correlation and T-Test Results—Bivariate correlations examined overall relationships between CV_{total} and the performance measures and then total pumps and the performance measures. CV_{total} was significantly negatively correlated with explosions ($r = -0.74$, $p < 0.001$), money earned ($r = -0.71$, $p < 0.001$), and total pumps ($r = -0.77$, $p < 0.001$). Total pumps was significantly positively associated with explosions ($r = 0.94$, $p < 0.001$) and with

money earned ($r = 0.55, p < 0.001$). Higher variability in pumps on the BART was associated with fewer explosions and fewer pumps, but also with poor performance in terms of money earned on the task. Using the standard assessment of risk-taking (total pumps), more risk-taking is associated with more explosions but more money earned.

A scatter-plot of CV and pumps revealed that participants with the highest CV scores also had the fewest total pumps. Mean comparisons indicated that participants with high CV scores (n had significantly fewer total pumps than low CV participants ($n = 7$) ($t(12) = 6.31, p < 0.001$), less money earned ($t(12) = 3.91, p < 0.01$), and fewer explosions ($t(12) = 6.62, p < 0.001$). Examination of intra-individual trial data on participants with the highest CV scores indicated that, on many trials, they chose to administer fewer than 20 total pumps (see Figure 1). As a result, their range of pumps used per balloon (i.e., one participant's pumps ranged from 3 to 100 across the trials) included more trials with very few pumps utilized.

Multiple Regression Results—Multiple regression models were used to determine whether CV was a significant predictor of BART performance when modeled simultaneously with pumps. In the first model, total pumps, CV, and money earned were entered as predictors of total explosions, and sex was entered as a covariate. The whole model was significant ($F(4,56) = 335.59, p < 0.001$, Adjusted $R^2 = 0.96$) and all predictors were significant. More pumps significantly predicted more explosions ($b = 0.91, p < 0.001$), and both CV ($b = -0.32, p < 0.001$) and money earned ($b = -0.40, p < 0.001$) predicted fewer explosions. Sex was not significant ($b = 0.01, p = 0.76$). In the second model, total pumps, CV and explosions were entered as predictors of money earned, with sex as a covariate. The whole model was significant ($F(4,56) = 69.48, p < 0.001$, Adjusted $R^2 = 0.83$) and all predictors were significant, except for sex ($b = 0.01, p = 0.87$). More pumps significantly predicted more money earned ($b = 1.57, p < 0.001$), whereas CV ($b = -0.76, p < 0.001$) and explosions ($b = -1.71, p < 0.001$) predicted less money earned. To assess whether these results were overly impacted by multicollinearity, we examined VIF and tolerance statistics. All were within acceptable limits. We also re-ran all models with predictors entered step-wise. All results were the same. CV, therefore, functioned as a significant and unique predictor of BART performance, even when modeled with pumps.

In the final model, CV was entered as a predictor of the difference score of pumps used on exploded balloons and pumps used on the balloons immediately after an exploded balloon. Gender was entered as a covariate. The model was significant ($F(2,282) = 49.18, p < 0.001$, Adjusted $R^2 = 0.26$), and higher CV was associated with greater difference scores ($b = 0.51, p < 0.001$). Sex showed a trend toward significance; females had a trend for greater difference scores, but it was not a statistically significant effect ($b = -0.09, p = 0.07$). More variability predicted greater change in pumps after exploded balloons. Participants with higher variability were more reactive after experiencing a negative outcome. In conjunction with the findings that variability is associated with fewer total explosions and with having more trials with less than 20 pumps, participants with higher variability appear to have more sensitivity to negative outcomes on the BART, even at the expense of earning more money.

Patterns of Performance across BART Trials

Four repeated measures ANOVAs were used to examine the effect of time on BART explosions, money earned, CV, and pumps (see Table 1).

Explosions—*Total explosions by time-block* (e.g. first 10 trials, second 10 trials, third 10 trials) was entered as the repeated measures variable. There was an overall effect of time, such that total explosions changed over the course of the trials ($F(2, 110) = 3.34, p < 0.05, \eta^2_p = 0.06$). Examination of the means and standard deviations (see Table 1) indicated that there was a decrease in explosions in Block 3. Polynomial contrasts indicated, in support of this, that there was a significant linear trend of decreasing explosions over the trials of the BART ($F(1, 55) = 4.50, p < 0.05$). The quadratic trend was not significant ($p = 0.15$), indicating that the increase from Block 1 to Block 2 was not significant. Therefore, there was an overall linear, decreasing trend on explosions, indicating that participants performed better over time.

Money Earned—There was an overall effect of time, such that total money earned changed during the trials, $F(1, 114) = 15.18, p < 0.001, \eta^2_p = 0.21$. Examination of the means and standard deviations (see Table 1) indicated that there was a decrease in money between Block 1 and Block 2 but an increase in money in Block 3. Polynomial contrasts indicated the presence of a significant quadratic trend, $F(1, 57) = 17.98, p < 0.001$, reflecting the higher level of money earned during Block 3 than during the other two Blocks. Therefore, there was an upward curvilinear relationship between money earned and time (i.e. more earned in Block 1, less in Block 2, and more in Block 3).

Pumps—There was an overall effect of time, such that pumps changed during the course of the trials, $F(2, 114) = 3.85, p < 0.05, \eta^2_p = 0.24$. Examination of means and standard deviations (see Table 1) indicated a small decrease from Block 1 to Block 2 and an increase in pumps during Block 3. Polynomial contrast results revealed that the linear trend was not significant, $F(1, 57) = 3.70, p = 0.06$, but that the quadratic trend was significant, $F(1, 57) = 4.05, p < 0.05$. Therefore, the linear decrease in pumps from Block 1 to Block 2 was not a significant trend, and overall there was an upward trend of pumps over time.

CV—There was a significant overall effect of time, such that variability in responding changed during the trials, $F(2, 112) = 3.95, p < 0.05, \eta^2_p = 0.07$. Polynomial contrast results indicated that, in contrast to the trends for pumps and money, the linear trend was significant ($F(1, 56) = 4.26, p < 0.05$) but the quadratic trend was not significant, $F(1, 56) = 3.67, p = 0.06$. Examination of the means and standard deviations (see Table 1) revealed an increase in CV from Block 1 to Blocks 2 and little change to Block 3. Therefore, the trend was for participants to exhibit more variability in risk-taking during the course of the BART.

Relationship to Alcohol Use

Exploratory bivariate correlations examined relationships among alcohol use, controlled drinking self-efficacy, injunctive norms, CV and pumps (see Table 2). CV was negatively related to peak drinking quantity but was positively related to peak drinking frequency. CV was also related to increased self-efficacy to control drinking in social situations and in

response to negative affect and was associated with lower acceptability of drinking and drinking-related consequences (i.e. injunctive norms). In contrast, total pumps showed the opposite relationships. Specifically, total pumps was positively related to peak drinking quantity, lower self-efficacy to control drinking, and higher injunctive norms. Higher variability in BART performance, therefore, is associated with less risky drinking and better perceived ability to control drinking. More total pumps, however, was associated with higher peak consumption, less confidence in the ability to control drinking, and a higher perception that one's friends approve of drinking.

Discussion

This study sought to determine whether CV provided a different picture of risk-taking behavior than the traditional BART measure of risk-taking. Past findings from animal studies (Cf. Weber et al., 2004) indicated that CV could be a better predictor of risk-taking behavior, and an animal model of the BART had shown that CV was negatively associated with food earned under risky conditions (Jentsch et al., 2010). Consistent with these findings, our results indicate that the CV provides novel information about risk-taking behavior in humans.

Higher variability was negatively associated with total money earned, which parallels the findings of Jentsch and colleagues (2010). Coupled with our unexpected finding that higher variability was negatively associated with total explosions, high variability appears to provide evidence of a risk aversive pattern of behavior. Because higher variability also predicted more change in response to exploded balloons, it appears that those with more variable responding sacrificed money to avoid more consequences.

In contrast, pumps, a traditional BART measure, was positively correlated with money earned and explosions, indicating that more risk was associated with more consequences but also with more reward. Over the trials, participants also decreased explosions and increased money earned, indicating that some learning effects took place. CV, therefore, provides a novel way of looking at BART performance and suggests that higher levels of variability can be associated with better avoidance of consequences and may reflect more adaptive functioning in relationship to risky behaviors, such as alcohol use.

Of interest, variability in other domains has been shown to be adaptive. Consistent with this, our results indicate that those with more variable responding on the BART had more self-efficacy to control drinking in social situations and in response to negative affect. Higher variability was also associated with lower quantity of peak drinking, but more frequency of peak drinking, suggesting that these participants drink at lower levels more often. Given that higher variability was associated with BART consequence avoidance, one hypothesis is that CV may correspond to a tendency to change behavior in response to or to avoid alcohol-related consequences. This suggests the possibility that a high CV pattern of responding may actually be adaptive over the long term. However, it is important to remember that this sample of young adults all drank heavily on a regular basis. Thus, prospective research involving CV in high-risk cohorts is needed before concluding that CV is adaptive.

Our findings stand in contrast to the recent finding that non-treatment seeking adults (ages 21–65, mean age = 30.29) with more alcohol use disorder symptoms were more conservative on the BART, as measured by pumps on unexploded balloons (Ashenhurst et al., 2011), and that inter-trial variability (defined as variability on unexploded balloons divided by mean unexploded balloons) was unrelated to alcohol use. This difference could be the result of a combination of methodological differences (e.g. allowing participants to specify total pump values versus clicks for each pump; 30 trials versus 72 trials) and sample differences (e.g. young adults in a clinical study of naltrexone versus community non-treatment seekers; young adults (18–25) versus adults (21–65)). Our finding that pumps was unrelated to peak drinking quantity is noteworthy because it matches their finding that adjusted pumps was unrelated to alcohol use. Replication is needed before generalizing to other samples.

Our findings should be considered in light of study limitations. This research consists of a small sample of heavy-drinking young adults seeking participation in a pharmacotherapy treatment trial. Though it is important to note that a considerable amount of compensation was offered, participants were not necessarily highly motivated to change their drinking behavior. Future studies could consider whether these findings generalize to other samples of heavy drinking young adults. Our sample also had significantly more males than females, and therefore could only utilize sex as a covariate. There was a trend effect for females to have greater reactivity after a negative consequence (e.g. an exploded balloon). It is possible that a sample with a more equal distribution of males and females would have found a significant effect, and future research should investigate sex differences on this metric. Sex differences in performance on the BART have been found in previous studies (e.g. Lighthall, Mather, & Gorlick, 2009), but we are not aware of any prior studies examining sex differences in levels of variability on the BART. Because of our small sample size, correlational results are considered exploratory. Future research with larger samples should examine these relationships and utilize corrections for multiple comparisons. It is also important to consider whether there are task specific effects. Future studies could consider whether different versions provide different assessments of risk-taking behavior and whether they are differentially related to alcohol use. Our sample was also more educated than samples used in some previous studies on the BART (e.g. Bickel, Yi, Landes, Hill, & Baxter, 2011; Lejuez et al., 2002). Higher variability could be an adaptive function in a more educated sample, but might reflect less inhibitory control in patients with more psychopathology.

This study represents the first human study to find that variability on the BART reflects a consequence-avoidant pattern of risk-taking. Additionally, it is the first to show that BART variability is significantly related to lower peak drinking quantity, greater perceived ability to control drinking, and lower acceptability of drinking and drinking-related consequences (i.e. lower injunctive norms). Our findings are consistent with results from pre-clinical research indicating that CV can provide additional information about risk-taking behavior. This human study helps clarify what CV might indicate; specifically, CV does not appear to indicate poor top-down behavioral control, as had been hypothesized. The inclusion of the CV in future analyses of behavioral risk tasks and models of risk-taking behavior in humans may provide a more accurate and thorough understanding of human choice.

Acknowledgements

The project described was supported by Grant Numbers R01 AA016621 (SSO), K05 AA014715 (SSO), K01 AA019694 (RFL), K23 AA020000 (LMF), T32 AA015496 (KSD) from the National Institute on Alcohol Abuse and Alcoholism, a grant from ABRMF/the Foundation for Alcohol Research (RFL) and by the Connecticut Department of Mental Health and Addiction Services. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

Dr. Stephanie O'Malley declares the following: Dr. O'Malley has served as a consultant to or advisory board member for Pfizer, Alkermes, Arkeo Pharmaceuticals, and the Hazelden Foundation; she is a member of the American Society of Clinical Psychopharmacology's Alcohol Clinical Trials Initiative, which is supported by Abbott Laboratories, Eli Lilly & Company, Lundbeck, Pfizer and Ethypharm; and she has received study supplies from Pfizer and a contract from Eli Lilly as a study site for a multi-site trial.

References

- Ashenhurst JR, Jentsch JD, Ray LA. Risk-taking and alcohol use disorders symptomatology in a sample of problem drinkers. *Exp Clin Psychopharmacol*. 2011; 19:361–370. [PubMed: 21707191]
- Bellgrove MA, Hester R, Garavan R. The functional neuroanatomical correlates of response variability: Evidence from a response inhibition task. *Neuropsychologia*. 2004; 42:1910–1916. [PubMed: 15381021]
- Bickel WK, Yi R, Landes RD, Hill PF, Baxter C. Remember the future: Working memory training decreases delay discounting among stimulant addicts. *Biological Psychiatry*. 2011; 69(3):260–265. [PubMed: 20965498]
- Carey KB, Scott-Sheldon LAJ, Carey MP, DeMartini KS. Individual-level interventions to reduce college student drinking: a meta-analytic review. *Addict Behav*. 2007; 32:2469–2494. [PubMed: 17590277]
- Cavalca E, Kong G, Liss T, Reynolds EK, Schepis TS, Lejuez CW, Krishnan-Sarin S. A preliminary experimental investigation of peer influence on risk-taking among adolescent smokers and non-smokers. *Drug and Alcohol Dependence*. 2012; 129:163–166. [PubMed: 23131775]
- Collins CR, Parks GA, Marlatt GA. Social determinants of alcohol consumption: the effects of social interaction and model status on the self-administration of alcohol. *J Consult Clin Psychol*. 1985; 53:189–200. [PubMed: 3998247]
- Dekker JM, Crow RS, Folsom AR, Hannan PJ, Liao D, Swenne CA, Schouten EG. Low heart rate variability in a 2-minute rhythm strip predicts risk of coronary heart disease and mortality from several causes: the ARIC study. *Circulation*. 2000; 102:1239–1244. [PubMed: 10982537]
- DeMartini KS, Palmer RS, Leeman RF, Corbin WR, Toll BA, Fucito LM, O'Malley SS. Drinking less and drinking smarter: direct and indirect protective strategies in young adults. *Psychology of Addictive Behaviors*. 2013; 27(3):615–626. [PubMed: 23088406]
- Dick DM, Smith G, Olausson P, Mitchell SH, Leeman RF, O'Malley SS, Sher K. Understanding the construct of impulsivity and its relationship to alcohol use disorders. *Addict Biology*. 2010; 15:217–226.
- Furby L, Beyth-Marom R. Risk taking in adolescence: a decision-making perspective. *Developmental Review*. 1992; 12:1–44.
- Hayano J, Sakakibara Y, Yamada A. Accuracy of assessment of cardiac vagal tone by heart rate variability in normal subjects. *Am J Cardiol*. 1991; 67:199–204. [PubMed: 1987723]
- Ingjaldsson JT, Laberg JC, Thayer JF. Reduced heart rate variability in chronic alcohol abuse: relationship with negative mood, chronic thought suppression, and compulsive drinking. *Biol Psychiatry*. 2003; 54:1427–1436. [PubMed: 14675808]
- Jentsch JD, Woods JA, Groman SM, Seu E. Behavioral characteristics and neural mechanisms mediating performance in a rodent version of the Balloon Analogue Risk Task. *Neuropsychopharmacology*. 2010; 35:1979–1806.
- Keppel, G.; Wickens, TD. Design and analysis: a researcher's handbook. 4th ed.. Upper Saddle River, NJ: Pearson Prentice Hall;

- Larimer ME, Turner AP, Mallet KA, Geisner IM. Predicting drinking behavior and alcohol-related problems among fraternity and sorority members: examining the role of descriptive and injunctive norms. *Psychol Addict Behav.* 2004; 18:203–212. [PubMed: 15482075]
- Leigh BC. Peril, chance, adventure: concepts of risk, alcohol use and risky behavior in young adults. *Addiction.* 1999; 3:371–383. [PubMed: 10605866]
- Lejuez CW, Read JP, Kahler CW, Richards JB, Ramsey SE, Stuart GL, Strong DR, Brown RA. Evaluation of a behavioral measure of risk-taking: the Balloon Analogue Risk Task (BART). *J Exp Psychol Applied.* 2002; 8:75–84. [PubMed: 12075692]
- Muraven M. Building self-control strength: practicing self-control leads to improved self-control performance. *Journal of Experimental Social Psychology.* 2010; 46:465–486. [PubMed: 20401323]
- Neighbors C, O'Connor RM, Lewis MA, Chawla N, Lee CM, Fossos N. The relative impact of injunctive norms on college student drinking: the role of reference group. *Psychol Addict Behav.* 2008; 22:576–581. [PubMed: 19071984]
- Pleskac TJ, Wallsten TS, Wang P, Lejuez CW. Development of an automatic response mode to improve the clinical utility of sequential risk-taking tasks. *Experimental and Clinical Psychopharmacology.* 2008; 16:555–564. [PubMed: 19086776]
- Shafir S, Wiegmann D, Smith BH, Real LA. Risk-sensitive foraging: Choice behavior of honey bees in response to variability in volume of reward. *Animal Behavior.* 1999; 57:1055–1061.
- Sitharthan T, Job RFS, Kavanagh DJ, Sitharthan G, Hough M. Development of a Controlled Drinking Self-Efficacy scale and appraising its relation to alcohol dependence. *J Clin Psychol.* 2003; 59:351–362. [PubMed: 12579550]
- Sitharthan T, Kavanagh DJ. Role of self-efficacy in predicting outcomes from a programme for controlled drinking. *Drug Alcohol Depend.* 1991; 27:87–94. [PubMed: 2029863]
- Thayer JF, Lane RD. A model of neurovisceral integration in emotion regulation and dysregulation. *J Affect Disord.* 2000; 61:201–216. [PubMed: 11163422]
- Weber EU, Hsee CK. Models and mosaics: Investigating cross-cultural differences in risk perception and risk preference. *Psychonomic Bulletin & Review.* 1999; 6:611–617. [PubMed: 10682203]
- Weber EU, Shafir S, Blais AR. Predicting risk sensitivity in humans and lower animals: Risk as variance or coefficient of variation. *Psychol Rev.* 2004; 111:430–445. [PubMed: 15065916]
- Wechsler H, Nelson TF, Lee JE, Seibring M, Lewis C, Keeling RP. Perception and reality: a national evaluation of social norms marketing interventions to reduce college students' heavy alcohol use. *J Stud Alcohol.* 2003; 64:484–494. [PubMed: 12921190]

Figure 1.

Pumps across BART Trials for a High CV versus a Low CV Participant

Note. High CV Participant: CV = 0.68, Total Pumps = 749, Total Explosions = 7. Low CV Participant: CV = 0.16, Total Pumps = 1866, Total Explosions = 14.

* explosion for High CV participant; † explosion for Low CV participant

Table 1
Means and Standard Deviations of Risk-Taking, Performance, and Variability on the BART

	Time Block					
	1		2		3	
	Mean	SD	Mean	SD	Mean	SD
Explosions	4.82	1.09	4.74	1.56	4.35	1.36
Money	266.17	66.41	248.16	59.45	304.03	74.43
Pumps	570.95	144.32	561.55	139.90	600.95	148.83
CV	0.28	0.15	0.34	0.21	0.32	0.19
					0.34	0.18

Note. BART = Balloon Analogue Risk Task. CV = Coefficient of Variability.

Table 2

Correlations among Peak Drinking, Protective Behavioral Strategies, Norms, Controlled Drinking Self-Efficacy and Performance on the Balloon Analogue Risk Task

	BART Performance	
	CV	Total Pumps
Peak Drinking – 3 months	–0.31 [*]	0.23
Peak Drinking – Frequency	0.32 [*]	–0.27 [*]
Peak Drinking - Lifetime	–0.31 [*]	0.26 [*]
PBS – Direct Strategies	–0.03	0.12
PBS – Indirect Strategies	0.01	0.11
Descriptive Norms	0.13	–0.03
Injunctive Norms	–0.39 ^{**}	0.32 [*]
Controlled Drinking Self-Efficacy – Neg. Affect	0.31 [*]	–0.24
Controlled Drinking Self-Efficacy - Social	0.48 ^{***}	–0.37 ^{**}

Note. BART = Balloon Analogue Risk Task. CV = Coefficient of Variability. PBS = Protective Behavioral Strategies.

^{*}
 $p < 0.05$

^{**}
 $p < 0.01$

^{***}
 $p < 0.001$