Modeling Behavioral Reactivity to Losses and Rewards on the Balloon Analogue Risk Task (BART): Moderation by Alcohol Problem Severity

James R. Ashenhurst, Spencer Bujarski, J. David Jentsch, and Lara A. Ray
University of California, Los Angeles

The relationship between risk-taking behavior and substance dependence has proven to be complex, particularly when examining across participants expressing a range of substance use problem severity. While main indices of risk-taking in the Balloon Analogue Risk Task (BART) positively associate with problematic alcohol use in adolescent populations (e.g., MacPherson, Magidson, Reynolds, Kahler, & Lejuez, 2010), several studies have observed a negative relationship when examining behavior within adult substance using populations (Ashenhurst, Jentsch, & Ray, 2011; Campbell, Samartgis, & Crowe, 2013). To examine potential mechanisms that underlie this negative relationship, we implemented multilevel regression models on trial-by-trial BART data gathered from 295 adult problem drinkers. These models accounted for participant behavior on trials following balloon bursts or cash outs as indices of loss and reward reactivity, respectively, and included control variables including age, IQ, and individual delay discounting rate. Results revealed that individual trial pumping was significantly predicted by trial number, and by whether or not the previous trial was a big burst or a big cash out (i.e., large magnitude of potential gains) in a manner consistent with a “near-miss” effect. Furthermore, severity of alcohol problems moderated the effect of a previous trial big burst, but not of a big cash out, on subsequent trial behavior such that those with greater severity demonstrated relative insensitivity to this “near-miss” effect. These results extend previous studies suggesting that alcohol abusers are less risky on the BART by specifying a mechanism underlying this pattern, namely, diminished reactivity to large magnitude losses.

Keywords: BART, alcoholism severity, near miss, risk-taking, decision-making

Alcohol use disorders are complex and multidimensional and may be understood, in part, by examining individual differences in endophenotypes, which are defined as heritable, biologically regulated behaviors or biomarkers that associate with disorder liability (Ducci & Goldman, 2008; Goldman, Oroszi, & Ducci, 2005). Given that clinical criteria for alcohol (and other substance) dependence focus heavily on continued use of the substance despite knowledge of the occurrence and risks of adverse health, legal, or social outcomes (American Psychiatric Association, 2000), researchers have posited that propensity for risk-taking, driven in part by genetic factors, may enhance liability for problematic substance use (Kreek, Nielsen, Butelman, & LaForge, 2005). While debate exists regarding the precise definition of risk-taking propensity, we view it as a pattern of maladaptive choice behavior produced under conditions where there exists a potential for reward but also an unknown probability of negative outcomes. This definition is consistent with the idea that risk-taking propensity may be a contributing factor to substance dependence liability, because continued substance use represents a decision to engage in maladaptive risky behavior to obtain a desired reward.

One increasingly popular task used to assess risk-taking propensity is the Balloon Analogue Risk Task (BART), which, as a behavioral task, is not subject to self-report bias (Lejuez et al., 2002). In this task, participants inflate a virtual balloon with a small potential payout per pump. However, the balloon may burst at any time, resulting in a forfeiture of earned money for that trial. Risk-taking is thus indexed by increased reward seeking in the face of greater potential loss. Consistent with behavior in this task being a potential endophenotype, twin studies and rat breeding studies have demonstrated that risk-taking behavior measured by the BART is moderately heritable (Anokhin, Goloshevin, Grant, & Heath, 2009; Ashenhurst, Seaman, & Jentsch, 2012).

Research using this task to examine the role of risky decision-making in substance use and misuse has yielded a complex picture,
particularly when examining individuals expressing differences in substance use problem severity (Ashenhurst et al., 2011) and across stages of substance use involvement. In subclinical and adolescent populations, increased risk-taking in the BART is associated with increased problematic alcohol and drug use (Fernie, Cole, Goudie, & Field, 2010; Lejuez, Akin, Bornova, & Moolchan, 2005; Lejuez et al., 2002; MacPherson et al., 2010). However, our group has demonstrated that risk-taking in the BART is negatively correlated with the severity of clinical alcohol dependence as defined by Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM–IV) criteria in adults with a range of alcohol use problems (Ashenhurst et al., 2011). The direction of this relationship was recently replicated in a study comparing long-term alcohol users versus nonusers in the BART, whereby the alcohol users took less risk overall in the BART than the nonusers (Campbell, Samartgis & Crowe, 2013).

A similar negative relationship was also described in a sample of young tobacco smokers, whereby smokers were less risk-taking on the BART than nonsmokers in terms of trajectory of balloon pumping across the test session (Dean, Sugar, Hellemann, & London, 2011). In addition, among adult smokers, tobacco dependence was negatively correlated with risk-taking in the BART (Ryan, Mackillop, & Carpenter, 2013), replicating a negative relationship between dependence severity and risk-taking assessed in the BART. Effects are not uniform, however, because multiple studies comparing tobacco dependent versus nondependent or tobacco using versus nonusing control participants have failed to detect overall differences in risk-taking propensity in the BART (Acheson & de Wit, 2008; Galván et al., 2013). Together these studies suggest that behavioral patterns in the BART are associated with the severity of substance use problems, yet the specific mechanisms explaining these differences remain unclear.

Modeling of data from another commonly used task to examine risk-taking behavior, the Iowa Gambling Task (IGT; Bechara, Damasio, Damasio, & Anderson, 1994), has offered a potential mechanistic explanation for how heavy substance users may differ from healthy controls in risk-taking tasks. In the IGT, participants are allowed to choose from among four virtual card decks, each of which has (unknown) probabilities of obtaining reward but also some losses. The decks are set up such that two decks yield net positive outcomes, while two yield net losses. With sufficient sampling, healthy participants learn to select the advantageous decks more than the disadvantages ones. However, individuals with substance use disorders tend to behave similarly to patients with prefrontal cortex damage; they fail to learn to select the advantageous decks, and demonstrate enhanced reward reactivity as well as diminished loss sensitivity, resulting in a risk-taking temperament with reduced economic utility in the IGT (Bechara & Damasio, 2002; Bechara et al., 2001; Stout, Busseymeyer, Lin, Grant, & Bonson, 2004).

Thus, the primary aim of this study was to implement a detailed trial-by-trial analysis of behavior in the BART to identify behavioral mechanisms that may explain why participants at greater levels of alcohol problem severity take less risk in the BART (Ashenhurst et al., 2011). In most analyses of the BART, data is tabulated as means across trials (calculating mean pumps on non-burst trials for a given administration of the BART), which fails to capture the trial-by-trial reactivity and learning that is occurring during the progression of the task.

Indeed, powerful novel multilevel regression-based approaches to individual trial data have yielded interesting results associating dopamine transporter gene variation with behavior in the BART (Mata, Hau, Papassotriopoulou, & Hertwig, 2012). Our first study aim was to replicate this model but instead of examining dopamine genetics, we examined whether alcohol problem severity moderates reactivity to rewards and losses on the BART. Considering findings from the IGT (Stout, Busseymeyer, Lin, Grant, & Bonson, 2004), our initial hypothesis was that individuals with more alcohol-related problems would be less loss reactive (i.e., modulate their behavior less following balloon burst trials) and more reactive to cash outs (i.e., increase pumping more following cash outs trials) than those reporting less severe alcohol problems.

The second study goal is to expand upon this modeling of trial-by-trial reactivity on the BART by accounting for both reward and loss magnitude to see whether sizes of the previous gambles (e.g., bursts that resulted in a large forfeiture) influenced participant behavior. Data from the gambling literature using slot machine-like tasks has demonstrated differences between reactivity to nearly won losses and clearly lost losses described as a “near miss” effect (Reid, 1986); in particular, participants tend to increase their willingness to gamble on trials following a “near miss,” defined as a failure that comes close to being highly successful. Thus, we sought to test whether the influence of reward and loss magnitude was moderated by alcohol problem severity.

The third study aim was to account for previously identified variables that influence performance of the BART, thus ruling out potential confounds. These control variables include demographic and neurocognitive indicators such as age, IQ, and working memory span (Ashenhurst et al., 2011). Furthermore, Dean et al. (2011) have suggested that the negative relationship between substance use and risk-taking in the BART may be because of the risk-taking being confounded with delay discounting because greater amounts of reward-seeking on trials in the BART (the outcome typically associated with more risky decisions) requires persistence and patience for a future (often small) reward. Consistent with this view, our group has shown, using a structural equation modeling approach, that risk-taking in the BART is negatively associated with delay discounting rates among problem drinkers. It is important that performance on the BART and the Delay Discounting Task were both related to alcohol problems, but in opposing directions (Courtney et al., 2012); risk-taking in the BART was negatively associated with greater alcohol problems, while delay discounting had a positive relationship. Thus, we included individual delay discounting rates as an additional control.

This study is the first, to our knowledge, to model both loss and reward reactivity on the BART directly in a population with substance use problems. In addition, this represents the first attempt to test whether severity of alcohol problems is associated with behavioral reactivity to bursts and cash outs of differing reward/loss magnitudes in the BART as assessed in a large community sample of heavy drinkers expressing a range of alcohol use problems. These analyses advance the literature by examining behavioral mechanistic explanations for the negative relationship between substance use problems and risk-taking in the BART in adults (Ashenhurst et al., 2011; Campbell et al., 2013; Ryan et al., 2013).
Method

Participants

Nontreatment seeking heavy drinkers \((N = 295)\) were recruited from the Los Angeles community through flyers, print, and online advertisements as part of a larger alcohol administration study. A subset of these participants \((N = 158)\) were included in a previous report on these data (Ashenhurst et al., 2011). Inclusion criteria were as follows: (1) age between 21 and 65; (2) self-identification of “problems with alcohol”; (3) telephone endorsement of consuming a minimum of 48 standard drinks per month. Exclusion criteria were as follows: (1) current treatment for alcohol problems, history of treatment in the 30 days prior to enrollment, or currently seeking treatment; (2) not having an alcoholic drink within 21 days of the telephone screening interview; (3) history of bipolar disorder or psychotic disorder, or positive evaluation for these disorders during a Structured Clinical Interview for DSM-IV (SCID; First, Spitzer, Gibbon, & Williams, 1995). Participants were compensated $40 for research participation, as well as up to an additional $5 based on performance on the BART (outlined in the Task Description section). The average age of the sample was 30.78 \((SD = 10.31, \text{range} = 21–63)\), with a majority of participants being male (73.14%). The ethnic background of the sample was as follows: White (42.9%), African American (18.8%), Asian (5.9%), Latino (13.5%), Native American (1.4%), Other/Mixed-Ethnicity (15.3%). The average number of years of education was 14.0 \((SD = 3.92)\). Descriptive statistics for alcohol use/problem indicators are presented in Table 1.

Procedures

Eligible nontreatment seeking individuals were invited to the laboratory for an in-person evaluation session, which included the following: the BART, the individual differences and alcohol problem severity measures, and a structured diagnostic interview. All participants provided written informed consent upon receiving a complete explanation of the study. Participants were required to have a blood alcohol concentration (BAC) equal to 0.000g/dl, as verified by a Breathalyzer test (Dräger, Telford, PA), prior to the testing session. All procedures were approved by the Institutional Review Board of the University of California, Los Angeles.

Measures

Participants were given a demographic questionnaire including: age, sex, ethnicity, and education. Additional study measures included those used to compute an alcohol problem severity score, the BART, and the Delay Discounting Task.

Alcohol problem severity. Severity of alcohol problems was indexed via a principle component score capturing a number of alcohol problems scales described by our group previously (Moalem, Courtney, Bacio, & Ray, 2013; Ray et al., 2013). Components included in the severity factor were the following. We used the SCID (First et al., 1995) to identify symptoms of alcohol abuse and alcohol dependence. These were recorded for a total of 11 possible symptoms (4 of abuse and 7 of dependence). Alcohol withdrawal was assessed using the Clinical Institute Withdrawal Assessment–Alcohol Revised (Puz & Stokes, 2005). The Penn Alcohol Craving Scale (PACS) captured craving for alcohol during the previous week (Flannery, Volpicelli, & Pettinati, 1999). A total score was also calculated from the Alcohol Dependence Scale (ADS; Skinner & Allen, 1982), a 25-item scale that measures alcohol dependence symptoms over the past 12 months. The Drinker Inventory of Consequences (DrInC-2R; Miller, Tonigan, & Longabaugh, 1995) provided a baseline description of the number and frequency of various drinking consequences, which was summed into a single indicator of negative drinking consequences. Under a principal component analysis, these five indicator variables comprised a single meaningful factor that explained 55% of the variance in alcohol problem indicators, with each indicator loading on the single factor >.40.

In previous analyses from our research group, this alcohol problem severity construct was related to subjective response to alcohol in an alcohol challenge (Ray et al., 2013), to affective symptoms and motivation to change (Moalem et al., 2013), and with fronto-striatal functional connectivity during performance of the Stop Signal Task (Courtney, Ghahremani, & Ray, 2013). For the present analyses, alcohol problem severity factor scores were centered and normalized to the sample.

The BART. A modified version of the BART (Lejuez et al., 2002) was administered as described previously (Ashenhurst et al., 2011). Briefly, participants were allowed to pump a virtual on-screen balloon and earn a small amount of money ($0.003) for each “pump”; these rewards are tailored continuously. At any point, the participant may stop pumping, add the earned reward to a guaranteed bank, and proceed to the next trial (a “cash out”). However, a certain level of risk was applied such that additional “pumping” might result in an on-screen burst of the balloon and a forfeiture of money earned for that one “burst” trial. Risk of balloon burst was distributed following a normal distribution with a mean at the midpoint of possible pumps (32 of 64 possible pumps) and a standard deviation of 20. At the end of 72 trials, accumulated earned totals were rounded to the nearest dollar for compensation. We chose this compensation rate as only about 12 minutes of the 3-hr visit were spent on the BART. Instead of collapsing all trial-by-trial data into single outcome measures as used previously (Ashenhurst et al., 2011; Lejuez et al., 2002), behavior on each trial was tallied for each participant, and entered into a multilevel regression model similar to one published previously (Mata et al., 2012) as described in data analysis.

Table 1

Means and SDs for the 5 Indicators of Alcohol Problem Severity

<table>
<thead>
<tr>
<th>Measure</th>
<th>(M)</th>
<th>(SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADS</td>
<td>40.25</td>
<td>7.31</td>
</tr>
<tr>
<td>PACS</td>
<td>17.92</td>
<td>6.61</td>
</tr>
<tr>
<td>Symptom count</td>
<td>5.23</td>
<td>2.81</td>
</tr>
<tr>
<td>DRINC-2R</td>
<td>40.9</td>
<td>22.12</td>
</tr>
<tr>
<td>CIWA-Ar</td>
<td>5.66</td>
<td>6.92</td>
</tr>
</tbody>
</table>

Note. ADS = Alcohol Dependence Scale; PACS = Penn Alcohol Craving Scale; symptom count = number of symptoms out of 11 (4 abuse and 7 dependence) from the Structured Clinical Interview for Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition axis I disorders (SCID-IV); DRINC-2R = Drinkers Inventory of Consequences-2 Revised; CIWA-Ar = Clinical Institute Withdrawal Assessment for Alcohol Revised.
Control variables. In a previous analysis from our group, we identified several neuropsychological variables that were related to performance in the BART and alcohol problems (Ashenhurst et al., 2011) including the following: (1) The Shipley Institute for Living Scale (Zachary, 1986) as an estimate of IQ; (2) The Digit Span Task as a classic working memory task that captures individuals’ abilities to cognitively retain and manipulate information. Norm-referenced scores from the Shipley scale and digit span task were used in the analyses as estimates of IQ and working memory, respectively.

Last, we included scores from the Delay Discounting Task to account for differences in temporal discounting. Participants were presented with a series of 27 hypothetical monetary choices and were asked to indicate their preferences between them. These choices were between small immediate rewards versus larger delayed rewards (e.g., $31 today, or $85 in 7 days). The parameters of these options were selected from a previously validated measure of delay discounting (Kirby, Petry, & Bickel, 1999). Participants were not compensated based on their choices, but were asked to consider them as real.

Delay discounting rates were computed by analyzing choice patterns fitted to the hyperbolic discounting functions derived from the following equation: \( V = A/(1 + kD) \), where \( V \) is the present value of the delayed reward \( A \) at delay \( D \), and \( k \) is a free parameter that determines the discount rate (Mazur, 1987). These \( k \) scores index the preference for smaller immediate rewards relative to larger delayed rewards. Three \( k \) variables were extracted from this measure, each pertaining to different magnitudes of reward: \( MS = $25, S55, $85; (1) k-Small, (2) k-Medium, and (3) k-Large, respectively. The average of these was computed as \( k-Total \) and used as a control variable in all analyses (log-transformed for normality considerations).

Data Analytic Plan

Analyses were conducted using a multilevel regression-based framework (Singer, 1998) using Proc Mixed in SAS version 9.3 for Windows. For all analyses, trials (Level 1) were nested within subjects (Level 2), and number of pumps on a given trial was the outcome variable. In the first set of models examining the effect of previous trial burst, Level 1 effects included the following: trial number (Trial; Level 1; coded 2 to 72, Trial 1 was excluded because data on previous trial could not be obtained), whether the current trial was a burst trial or not (Burst; coded success = 0, burst = 1), and whether the previous trial was a burst trial or not (Burst\(_{\text{-1}}\); coded previous success = 0, previous burst = 1). Level 1 effects were treated as random effects at the subject level with an unstructured variance/covariance matrix and Satterthwaite approximated degrees of freedom. Approximately as third of the variance in pumps was between-subjects variance (intraclass correlation [ICC] = 0.35), necessitating a multilevel nested approach to these analyses. The following set of equations was used to model behavior on the BART based on previously published models (Mata et al., 2012):

In Level 1, pumps on a given trial is predicted by a linear combination of the intercept (\( \beta_{0i} \)), the trial number (\( \beta_{1i} \)), whether that trial was a burst trial (\( \beta_{2i} \)), whether the previous trial was a burst trial, indexing reactivity to bursts (\( \beta_{3i} \)). These estimates then serve as outcomes at the subject level (Level 2), where severity of alcohol problems was allowed to predict intercept (simple effect of Severity, \( \gamma_{0i} \)) as well as reactivity to Bursts (Severity \( \times \) Burst\(_{\text{-1}}\), \( \gamma_{1i} \)), capturing whether alcohol problem severity moderated reactivity to Bursts.

Level 1: Pumps = \( \beta_{0i} + \beta_{1i} \) (Trial) + \( \beta_{2i} \) (Burst\(_i\))
+ \( \beta_{3i} \) (Burst\(_{\text{-1}}\)) + \( \epsilon_{ii} \)

Level 2: \( \beta_{0i} = \gamma_{00} + \gamma_{01} \) (Severity) + \( \gamma_{0j} \) (Covariates) + \( u_{0i} \)
\( \beta_{1i} = \gamma_{10} + u_{1i} \)
\( \beta_{2i} = \gamma_{20} + u_{2i} \)
\( \beta_{3i} = \gamma_{30} + \gamma_{3j} \) (Severity) + \( u_{3i} \)

* For ease of presentation, covariates (e.g., age, ethnicity, and IQ) are represented as a single variable.

In the second set of models examining the effect of cash out and burst magnitude, Level 1 effects included the following: Trial, Burst\(_i\), previous trial was a big burst (Big Burst\(_{\text{-1}}\); coded 0–75th percentile pumped burst trial for a given participant = 0, top 25th percentile pumped burst trial = 1), whether the previous trial was a typical burst (Typical Burst\(_{\text{-1}}\); reverse coded of Big Burst\(_{\text{-1}}\)), and whether the previous trial was a big cash out (Big Cash Out\(_{\text{-1}}\); coded 0–75th percentile pumped cash-out trial for a given participant = 0, top 25th percentile cash-out trials = 1). In this coding scheme, a typical cash-out trial (i.e., bottom 75th percentile of cash-out trials) was the reference group. Given the added complexity of this model, a hierarchical modeling approach was employed where Level 1 effects were entered in Block 1 and in subsequent blocks, subject-level variables of interest were entered as both main effects and moderators of response to previous trial characteristics (e.g., Severity \( \times \) Big Burst\(_{\text{-1}}\); Table 2). Again, all Level 1 effects were treated as random at the subject level with an unstructured variance/covariance matrix.

Results

Baseline Characteristics and Alcohol Problem Severity

Participants exhibited a wide range of alcohol problems as indicated by incidence of clinical criteria from the diagnostic interview. Specifically, a majority met criteria for alcohol dependence (72%), while a subset met criteria for abuse only (12%) or did not meet criteria for either abuse or dependence (16%). Descriptive statistics for the other alcohol problem measures used to compute alcohol problem severity factor scores are presented in Table 1.

Participants pumped on average 17.57 times per trial (SD = 5.95), and 23% of trials burst. On average, session totals were $3.44 (SD = $0.51).

Burst Reactivity

In a main effects only model, there was a significant effect of trial number (\( \beta = -0.017, SE = 0.003, p < .0001 \)) after controlling for age (\( \beta = -0.01, SE = 0.02, p = .08 \)), ethnicity (p value range: 0.01 – 0.48), working memory (\( \beta = -0.06, SE = 0.08, p = .47 \)) and IQ (\( \beta = 0.014, SE = 0.013, p = .90 \)). A significant main effect of Burst was observed (\( \beta = 0.39, SE = 0.19, p < .05 \)) as was a significant main effect of previous trial burst (Burst\(_{\text{-1}}\), \( \beta = 1.26, SE = 0.12 \),...
A significant main effect of alcohol problem severity (β = −0.43, SE = 0.19, p < .05) was observed. In a subsequent model a significant Severity × Big Burst−1 interaction was observed (β = −0.22, SE = 0.11, p < 0.05) such that as Severity increased, participants pumped fewer times after a burst trial compared with a cash-out trial (Figure 1). This moderation effect was unaffected by removal of covariates (Severity × Burst−1 in a model without covariates: β = −0.20, SE = 0.10, p < .05).

Magnitude of Bursts and Rewards

To examine the influence of burst and reward magnitude, a series of multilevel models were conducted wherein the previous trial characteristic was coded as either a Big Burst, a Typical Burst, a Typical Cash Out (reference group) or a Big Cash Out (coding scheme above in Data Analytic Plan). Overall, 7% of trials were coded as Big Bursts, 17% as Typical Bursts, and 24% as Big Cash Outs, with the remaining 52% of trials serving as the Typical Cash Out reference group.

In Block 1 (i.e., Level 1 effects only) we observed significant main effects of Big Burst−1 (β = 0.91, SE = 0.17, p < .0001), Typical Burst−1 (β = −1.13, SE = 0.11, p < .0001), and Big Cash Out−1 (β = 2.39, SE = 0.12, p < .0001) such that pumping increased overall after big cash outs and big bursts and decreased after typical bursts (Table 2). Furthermore, severity of alcohol problems was found to moderate reactivity to big bursts only (Severity × Big Burst−1: β = −0.40, SE = 0.17, p < .05). This moderation was followed-up with a series of post hoc analyses with recentered models to determine regions of significance between trials following typical cash-outs (reference group) and Big Burst−1 trials. These post hoc analyses revealed significant (at p < .05) simple effects of Big Burst−1 among participants with ≥1.1 SDs above the mean on Severity, and no significant Big Burst−1 effect at 1.2 SD or more (Figure 2). Severity was not found to moderate response to big cash outs or to typical bursts (p > 0.29). As is shown in Table 2, the significance or magnitude of these effects was not significantly impacted by the inclusion of subject-level covariates.

BART Performance and Delay Discounting Rate

Controlling for delay discounting rate as both a main effect, as well as a moderator of reactivity to bursts and cash outs, did not substantively alter any of the results presented. Alone, delay discounting rate (log transformed) was significantly associated with number of pumps on the BART (β = −0.36, SE = 0.16, p < .05) in the hypothesized direction, however this effect was not robust to controlling for age (k-Total: β = −0.18, SE = 1.5, p = .23; Age: β = −0.11, SE = 0.02, p < .0001). Furthermore, k-Total was not found to moderate responses to burst or cash-out trials (either big or typical; all ps ≥ .15).

Sex Differences

Sex was investigated as a covariate to assess the generalizability of these findings. In terms of overall burst reactivity, no main
A greater severity predicts a greater burst reactivity. The parameterization of our first model closely follows one published previously (Mata et al., 2012). In this simpler model where magnitude of the gamble is not accounted for, our results show a similar effect of a previous burst trial to that of Mata et al. (2012); that is, on trials following a balloon burst, participants tended to take less risk. It is interesting that the effect of previous trial bursts was significantly moderated by alcohol problem severity (Figure 1), indicating that participants expressing more problems took less risk following a prior burst than participants with lesser problem severity. This model suggests that the more severe participants were less risk-taking in part because they were more reactive to recent failure.

Counter to data from previous modeling (Mata et al., 2012), there was a small but significant effect of trial number such that across the testing session, participants took less risk as trials progressed. This inconsistent result may be because of different numbers of trials between implementations of the task; Mata et al. (2012) used a 30 trial variant, while we used a version with 72 trials. Thus, we may have observed a small degree of participant fatigue. However, the magnitude of this effect was quite small,

effect of sex was found ($\beta = 0.08, SE = 0.36, p = .83$), and sex was not found to moderate reactivity to burst trials ($\beta = 0.23, SE = 0.23, p = .31$).

Furthermore, the inclusion of sex effects did not alter the significance nor magnitude of the observed Severity × Burst,−1 interaction and no Sex × Severity × Burst,−1 interaction was observed ($p = .26$). In models of burst and reward magnitude, no main effect of sex was observed ($\beta = 0.26, SE = 0.38, p = .49$); however, sex was found to moderate the effect of Big Cash Out,−1 ($\beta = -0.86, SE = 0.28, p < .01$). This moderation was such that the effect of Big Cash Out,−1 was greater for females compared with males ($\beta_s = 3.06, and 2.20$, respectively, $p < .0001$). Sex was not found to moderate responses to big or small bursts ($p < .49$), and controlling for sex both as a main effect and moderator did not alter the severity of alcohol effects in terms of magnitude or significance. Last, no Sex × Severity × previous trial characteristic interactions were observed ($ps \geq .27$). Thus, while sex differences were observed in terms of reward reactivity, their presence does not preclude interpretation of any of the primary results reported.

Discussion

The goal of these analyses was to identify behavioral mechanisms underlying the previously observed negative relationship between risk-taking in the BART and substance dependence (Ashenhurst et al., 2011; Dean et al., 2011; Ryan et al., 2013). This was accomplished by implementing a multilevel regression model to examine trial-by-trial behavior while taking into account behavioral reactivity to bursts (failures) and cash outs (rewards). In addition, we sought to categorize such trials by magnitude of the gamble at stake and to control for important demographic and neurocognitive variables such as age, IQ, and individual delay discounting rate.

### Figure 1

Alcohol problem severity moderates Balloon Analogue Risk Task (BART) reactivity. Predicted values are from a model without covariates and are shown for Trial No. 32 (i.e., the middle of the task) and for nonburst trials (i.e., Burst, = 0). Overall, participants pumped less after a burst trial compared with after a cash-out trial ($p < .0001$). In addition, severity of alcohol problems was found to moderate reactivity to bursts ($p < .05$), such as level of alcoholism severity increased participants pumped fewer times after a burst trial compared with a cash-out trial.

### Figure 2

Predicted pump values based on a multilevel regression model examining the effect of alcohol problem severity on Balloon Analogue Risk Task (BART) behavior. Predicted values are from a model without covariates and are shown for Trial No. 32 (i.e., the middle of the task) and for nonburst trials (i.e., Burst, = 0). Overall, previous trial characteristics were highly influential on BART behavior, such that pumping increased after both big cash outs and big bursts and decreased after a typical burst (all $p < 0.0001$). In addition, severity of alcohol problems was found to moderate reactivity to big bursts ($p < .05$), such that increased pumping following a big burst (i.e., the "near-miss" effect) was only seen at low levels of alcohol problem severity. Alcohol problem severity did not moderate response to big cash outs or typical bursts. Post hoc analyses are represented such that a dotted line indicates no significant simple effect of Big Burst,−1. Post hoc analyses showed that among the most severe 42 subjects (Severity $\geq 1.1$ SD above the mean), no significant "near-miss" effect was observed.
suggesting a predicted decrease in pumps across the session on the order of about less than two pumps.

**Magnitude of Gambles and Alcohol Problem Severity**

The extension of our primary aim of this analysis was to assess whether alcohol problem severity was related to both burst and cash out reactivity taking into account the magnitude of the gamble in the previous trial. This analysis is the first, to our knowledge, to account for reactivity to both bursts and cash outs of differing magnitudes in a trial-by-trial analysis.

Consistent with our expectations, overall, participants took less risk on trials following typical balloon “bursts” (trial failures) and took more risk on trials following big cash outs (in the top 25th percentile; Figure 2). Intriguingly, participants tended to pump more on trials following big bursts (top 25th percentile of potential earnings) compared with typical cash outs. This observation is consistent with theory from the problem gambling literature on a “near miss” effect (e.g., Reid, 1986). A “near miss” is defined as a failure that comes close to being highly successful. Trials where participants successfully pumped to larger magnitudes but then were faced with a balloon burst may have been perceived as near misses. Near misses have been shown to increase motivation to voluntarily spend more time gambling and to bet more money in slot machine-like tasks (Côté, Caron, Aubert, Desrochers, & Ladouceur, 2004). A neuroimaging study indicated that neural responses to near misses in the striatum and the insula were similar to responses to wins, which may drive an increase in subsequent gambling despite the lack of actual reward delivery (Clark, Lawrence, Astley-Jones, & Gray, 2009). This activity may contribute to pathological gambling, as participants expressing greater severity of gambling problems show greater ventral striatal response to near misses (Chase & Clark, 2010). Although the BART was not designed to identify this effect, it is plausible that these near miss-like experiences in big burst trials encouraged greater levels of pumping on the proceeding trial.

Results taking into account alcohol problem severity, however, were partially consistent with our initial hypotheses that greater alcohol problem severity would predict enhanced reward but reduced loss reactivity as has been demonstrated in analysis of the IGT (Stout et al., 2004). Our results demonstrate that alcohol problem severity moderated reactivity to losses, but not to rewards. Specifically, participants with greater alcohol problem severity were less subject to a “near miss” effect than participants with less severe alcohol problems; these severe participants did not increase their pumping after experiencing a big burst compared with a typical cash out such that among the most severe 42 subjects (Severity > 1.1 SD above the mean), no significant “near miss” effect was observed (Figure 2). However, while this post hoc strategy implies a step-like effect of Severity (i.e., “near miss” effect in most subjects, no “near miss” effect among only the most severe), in our analyses severity of alcohol problems was analyzed as continuous covariate and thus our results suggest that as severity increased across the entire range of alcoholism severity the magnitude of an apparent “near miss” effect decreased.

Thus, these more severe participants did indeed demonstrate a blunted response to bursts, but only after trials with large gambles at stake. On the other hand, alcohol problem severity did not modulate the difference between reactivity to big cash outs versus typical cash outs. These results provide one potential mechanistic explanation for why greater alcohol problems are associated with less overall risk-taking in the BART (Ashenhurst et al., 2011), namely, a blunted near miss-like effect among those with greater alcohol problems. Lastly, Cook’s D values were computed for each subject to assess whether individual cases were driving the effects reported. While several cases were found to have high influence relative to the overall sample (e.g., >0.03), removal of any or all of these cases from the burst reactivity and burst/cash out magnitude models did not alter the significance of the substantive findings.

**Comparing the BART With Other Risk-Taking Tasks**

Our primary model showed decreased pumping after burst trials, and this effect appeared to be more robust in those with more severe alcohol problems. In subsequent models examining the impact of reward and loss magnitude, our results suggested that, while severity of alcohol problems did not moderate behavioral response to typical bursts, it was found to moderate response to big bursts. Thus, contrary to the tentative conclusions one would draw from the first model, namely that more participants with more alcohol problems were more responsive to losses, results from the second set of models demonstrate that severity of alcohol problems was negatively associated with magnitude of the near miss effect. This more specific effect related to large magnitude losses, then, explains the moderated relationship observed in the first set of models.

These findings stand in partial contrast to models of behavior in the IGT, where substance users are found to be more reward sensitive and less loss reactive. The gap between the findings with the IGT and ours with the BART may be because of significant differences between the tasks, the populations studied, as well as the methods for analyzing behavior. While both tasks require sampling and learning to improve performance, the nature of optimal behavior does differ between them. In the IGT, less risk-taking is always a more advantageous choice, while in the BART, less risk-taking actually results in reduced economic utility; this is because optimum performance in the BART involves balancing an increase in reward with an increase in risk, resulting in a nonlinear function (Jentsch, Woods, Groman, & Seu, 2010; Lejuez et al., 2002). Next, our analysis is within a substance abusing population and does not compare abusers to healthy controls. Finally, we allowed loss and reward reactivity to operate independently in our statistical model, while models of the IGT restrict these two factors to being on one dimension represented by a single parameter (Stout et al., 2004).

While recent coadministration of these two tasks in healthy controls demonstrated a positive relationship between risk-taking indicators in these tasks, this relationship was only significant in repeated administrations of the task, not among task-naïve participants (Xu, Korczykowski, Zhu, & Rao, 2013). While both the IGT and the BART are kinds of risk-taking tasks, key differences in task structure may recruit decision-related circuitry (that may be impacted by alcohol misuse) in different ways. Future research into task-specific neurocognitive deficits or biases caused by or antecedent to problematic alcohol use may enhance our understanding of the neurobehavioral architecture of alcohol and other substance use disorders. Future studies should coadminister the
BART, IGT, and other risk-taking tasks in clinical populations to evaluate the cross-task validity of risk-taking indicators because specific aspects of task design may subtly influence behavior.

Our results should be weighed with respect to the strengths and limitations of this study design. Our strengths included assessment of a large community sample of problem drinkers, extension of a previously published novel method to examine trial-by-trial behavior in the BART (Mata et al., 2012), and controlling for demographic variables, general intelligence, and temporal discounting rates in all analyses. Furthermore, while many studies compare groups of dependents versus nondependents or users versus nonusers (e.g., Campbell et al., 2013; Dean et al., 2011; Stout et al., 2004), examining within a sample expressing a broad range of alcohol use problems reveals the nature of continuous relationships between these constructs.

Limitations of the study included a somewhat restricted range of alcohol problem severity, as our sample was comprised of individuals who self-identify as having “problems with alcohol” and had to meet a minimum alcohol consumption level. Thus, neither nondrinking controls nor social drinkers who do not identify as having alcohol problems were included.

Future research comparing between users and nonusers is warranted. Nevertheless, there was a broad range of alcohol problems represented, as a substantial minority did not meet clinical criteria (DSM-IV-TR) for either dependence or abuse (16%). Next, this study was implemented as a cross-sectional design, which precludes causal inferences; it is unclear whether the moderating effect of alcohol problem severity on large magnitude loss reactivity is either a cause or consequence of problematic alcohol use.

Last, the version of the BART used here differs from prior implementations in several ways, potentially imposing some limits on the generalizability of these data. The payout schedule is lower compared with most versions of the task (e.g., Dean et al., 2011; Lejuez et al., 2002; Mata et al., 2012). It is important that data indicates that the value of a balloon pump does affect decision biases in the BART, with those with low impulsivity/sensation seeking self-report scores being the most sensitive; specifically, as the value of a single pump increases (e.g., from one penny to one quarter per pump), participants are generally less willing to take more risk (Acheson & de Wit, 2008; Bornovalova et al., 2009). Because the payout schedule implemented here is lower than most versions, this suggests that participants may have taken more risk here than under typical implementations of the task, and that any moderating effects of individual differences in impulsivity/sensation seeking may have been reduced. Second, the risk function differed in terms of the riskiness applied to a given balloon. The version presented here guarantees a balloon burst within 64 pumps, while prior versions have used a range of up to 8, 32, or 128 pumps (Lejuez et al., 2002) with only the last risk version (128 pump range) significantly correlating with self-reported real-world risk behavior. Indeed, most studies have used this risk level (e.g., Dean et al., 2011; MacPherson et al., 2010; Mata et al., 2012), potentially impacting the strength of the relationship between behavior reported here and real-world manifestations of risk-taking.

Conclusions

In sum, this study examined a multilevel regression analyses of trial-by-trial behavior in the BART (Mata et al., 2012) in a large sample of adults with a range of alcohol problem severity. We observed that participants with greater alcohol problem severity were less risk-taking in the face of a recent burst trial than participants with lesser severity. We extended this initial model by including parameterization of magnitudes of both cash out and burst trials to gauge behavioral reactivity in the BART. We found that with greater alcohol problem severity, participants were less subject to a “near miss” effect, providing a more specific account for why more clinically severe participants take less risk, overall, in the BART (Acheson et al., 2011; Ryan et al., 2013). In addition, we confirmed a negative relationship between delay discounting and BART risk-taking as shown previously in SEM modeling (Courtney et al., 2012) by using a different hierarchical regression approach and trial-by-trial modeling. Critically, our analyses survived controlling for previously implicated and theoretically important covariates of BART performance including delay discounting rate, IQ, and working memory span.

As others have observed a negative relationship between risk-taking in the BART and substance dependence in adult tobacco and alcohol users (Campbell et al., 2013; Ryan et al., 2013), future studies should examine trial-by-trial behavior to more fully evaluate behavior in the task. Decision-making under risk represents a complex cognitive process that is likely influenced by subteltes of task design. Still, observed differences in behavior within clinical populations and between substance dependent individuals and healthy controls are likely to indicate neurocognitive factors that partially explain liability to problematic substance use.

References


MODELING REACTIVITY IN THE BART


Received November 5, 2013
Revision received March 10, 2014
Accepted April 1, 2014