A Daily Measure of Positive and Negative Alcohol Expectancies and Evaluations: Documenting a Two-Factor Structure and Within- and Between-Person Variability

CHRISTINE M. LEE, PH.D.,[a],* DAVID C. ATKINS, PH.D.,[a] JESSICA M. CRONCE, PH.D.,[a] THERESA WALTER,[a] AND BARBARA C. LEIGH, PH.D.[b]

[a]Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, Washington
[b]Alcohol and Drug Abuse Institute, University of Washington, Seattle, Washington

ABSTRACT. **Objective:** Despite the documented importance of alcohol outcome expectancies in predicting alcohol use and related consequences, little research has explored within-person variability in expectancies. This article details the construction and psychometric analysis of a measure of alcohol expectancies specifically designed for daily assessment. **Method:** We developed a 15-item instrument to measure the likelihood of experiencing various outcomes from drinking, as well as the subjective evaluation of these outcomes. College students ($N = 352$; mean age = 19.7 years, $SD = 1.26$; 53.4% female) participated in a yearlong study wherein they completed three computerized interviews daily via mobile phones for 2 weeks in each academic quarter. Multilevel exploratory factor analysis was used to examine dimensionality at between-person and within-person levels, and generalizability coefficients were calculated to establish reliability. **Results:** Intraclass correlation coefficients were generally between .30 and .40, demonstrating both between-person and within-person variability. Exploratory factor analysis resulted in a two-factor solution of positive and negative effects of alcohol, and two items with equivocal loadings were dropped from the final scale. The two subscales showed excellent reliabilities at within- and between-person levels, and the measure demonstrated good convergent and discriminant validity with a commonly used expectancy measure. **Conclusions:** Drinkers hold many expectations about the effects of alcohol, and measures are needed that are designed to capture both stable and context-dependent aspects of these beliefs. Results demonstrated significant day-to-day variation in the strength and valuation of alcohol expectancies, and the scale demonstrated good psychometric properties that establish its appropriateness for use in daily process studies of alcohol use. (*J. Stud. Alcohol Drugs, 76,* 326–335, 2015)

ALCOHOL REMAINS THE SUBSTANCE MOST widely used by college students, with 67.7% reporting any alcohol consumption within the past 30 days, 40.1% reporting having "been drunk" during this same time period, and 37.4% reporting heavy episodic consumption (five or more drinks in a row) at least once in the past 2 weeks (Johnston et al., 2013). Alcohol use can be associated with negative consequences, ranging from academic impairment and sleep/mood disturbance to injury, legal involvement, or death (Perkins, 2002), and the risk for negative consequences increases exponentially with the number of heavy drinking episodes (Wechsler et al., 1998).

Various potential etiological factors of high-risk drinking among college students have been explored, including alcohol outcome expectancies. Expectancies are beliefs about positive and negative consequences that will result from drinking—specifically, effects on one's behavior or experience, such as increased sociability or tension reduction. Alcohol outcome expectancies reliably predict both frequency and quantity of alcohol consumption as well as risk for alcohol-related harm (see Jones et al., 2001, for a review). Most studies of alcohol expectancies have treated these cognitions as a stable trait of the individual (e.g., Brown et al., 1987; Fromme et al., 1993) rather than a fluid product of past and present experience, even when examining daily process models of alcohol consumption (e.g., Armeli et al., 2000, 2007). By comparison, consistent with what would be expected from a reciprocal model, some daily process studies have shown that certain expectancies and the evaluations of some alcohol outcomes vary significantly within persons across time (Armeli et al., 2005; Butler et al., 2010). Thus, research suggests that there are both trait and state qualities of alcohol expectancies, but scale development work supporting these stable and context-dependent aspects is limited.

As more research is moving toward examining repeated measures of alcohol expectancies, it is important to develop a scale of alcohol expectancies that can capture specific expectancies on a particular day, occasion, or event. A major advantage of multiple reports of expectancies (e.g., daily reports) is the facility to examine the relationships between events and their precursors at the level of the event itself. For example, using daily reports to study these links ensures the temporal contiguity and ordering of the events and their precursors and can reduce memory biases of retrospective

interviews. A number of studies are increasingly using multiple repeated measurements (e.g., daily, weekly) of the precursors (e.g., daily mood, daily stress, situational contexts; for example, Helzer et al., 2006; Mohr et al., 2005) and consequences of drinking. However, researchers need to consider whether the instruments used to measure psychological constructs (e.g., alcohol expectancies) are appropriate for daily or repeated use. In existing instruments that measure alcohol expectancies, respondents answer questions about the likelihood that drinking would result in various effects in unspecified situations, requiring respondents to abstract their behavior over a number of hypothetical situations rather than rating the likelihood for a specific time and place. The resulting answers depend on the inference rules (such as enumeration or rate-based strategies) that are used to construct the responses (Brown, 1995; Burton & Blair, 1991).

Existing alcohol expectancy instruments thus treat alcohol expectancies as a stable construct, or a trait rather than a state, and the resulting score is a summary of expectancies aggregated over all drinking situations. Therefore, researchers cannot be sure what the answers mean for specific situations. Expectancy ratings can be different depending on the amount drunk, location, and activities (Ham et al., 2013; Krank et al., 2005; MacLatchy-Gaudet & Stewart, 2001; Wall et al., 2000, 2001, 2003). Thus, an expectancy measure that is used in studies that incorporate repeated measures of drinking and consequences must capture what respondents expect from drinking on particular occasions. Moreover, several issues should be considered when developing a measure of alcohol expectancies to be used in daily process studies.

Brevity

In a study that asks people to answer the same questions repeatedly, issues of respondent motivation and compliance are paramount. A measure of alcohol expectancies used in this context needs to be as short as possible without sacrificing the needs of the research itself. "Short" incorporates several parameters: the number of questions, the length (number of words) of each question, and the number of words used in ancillary text such as introductions, instructions, and descriptions of outcome scales.

Positive and negative drinking outcomes

Although some alcohol expectancy measures focus solely on expectations of positive (Brown et al., 1980) or negative (Jones & McMahon, 2003) drinking outcomes, several others (Fromme et al., 1993; Leigh & Stacy, 1993) incorporate expectancies of both kinds of outcomes. Including a variety of both positive and negative expectancies enables the examination of the predictive effects of individual expectancies (e.g., feel more relaxed) and classes of expectancies (positive vs. negative).

Evaluation of consequences

Labeling alcohol effects as "positive" or "negative" can be problematic, given that effects that are desirable for some people may be undesirable for others (Leigh & Stacy, 1994; Mallett et al., 2008). Expectancy-value theories of attitudes and behavior (Ajzen & Fishbein, 1980) emphasize that the value attached to behavioral outcomes contributes to attitudes and intentions with respect to the behavior. The evaluations of particular drinking consequences might vary not only across people but also within people, and a repeated assessment of evaluations allows examination of this changing "payoff" of drinking for the individual in different situations.

Uncommon drinking outcomes

Certain drinking consequences, often the most severe, are rare. Consequences such as being arrested or severely injured are never experienced by most drinkers and are unlikely to happen to participants during a research study. Including these consequences in a survey measure that is administered daily may not be advisable. Asking participants repeatedly about the likelihood that these unlikely consequences will occur today can become annoying, and the answer shows no variability.

Purpose of the present study

This article describes the process of developing a daily measure of alcohol expectancies and evaluations of those expectancies as part of a larger program of research that examines a daily process model of alcohol use, alcohol expectancies, and alcohol-related consequences. In this article, we (a) describe scale properties, (b) demonstrate within- and between-person variability in alcohol expectancies and evaluations at the item and subscale level, and (c) establish basic convergent and discriminant validity with a common, cross-sectional assessment of expectancies. We hypothesized that the scale would have a reliable two-factor structure including both positive and negative expectancies and evaluations, and that items and scales would demonstrate both within- and between-person variability.

Method

Study 1: Item development and selection

The process of item development and selection began with the first and last authors reviewing existing alcohol outcome expectancy scales (including but not limited to the Comprehensive Effects of Alcohol measure [Fromme et al., 1993], the Alcohol Expectancy Questionnaire [Brown et al., 1987], and the National Alcohol Survey [Leigh & Stacy, 1993]) and constructing items for an expectancy measure

TABLE 1.   Daily expectancy and evaluation items

1. Feel more relaxed
2. Have a hangover
3. Be more sociable
4. Become aggressive
5. Be in a better mood
6. Feel nauseated or vomit
7. Hurt or injure yourself by accident
8. Get a buzz
9. Be unable to remember what you did while drinking
10. Have more desire for sex*
11. Be unable to study*
12. Feel more energetic
13. Be rude or obnoxious
14. Be able to express your feelings more easily
15. Do something that embarrassed you

Note: Asterisks (*) denote the two items that were removed from the final measures.

that incorporated a variety of domains of positive and negative alcohol effects. An initial set of 17 items was identified (including 9 negative and 8 positive items). We then used cognitive interviewing procedures, including verbal probing (Beatty & Willis, 2007; Willis, 2005), to evaluate comprehension and interpretation of instructions and scale items, revise the wording of instructions, and modify questionnaire items to include those most relevant to participants. Fourteen undergraduate students who typically consumed at least two drinks on one occasion each week were compensated $20 for their participation in the cognitive interviews.

Based on these interviews, slight revisions were made, which included dropping a few alcohol effects that showed no variability (these were consistently rated as always very bad and always very unlikely), revising the wording of questions that were occasionally misinterpreted (for example, adding the words "by accident" to the item "I would hurt or injure myself"), and combining redundant items. The final instrument (Table 1) incorporated 15 items.

Study 2: Daily administration

Participants. Participants were 352 undergraduate college students (M_{age} = 19.7 years, SD = 1.26; 53.4% female). Most participants (74.2%) were White, with the remainder Asian American (8.5%), multiracial (11.1%), or other (6.2%). Because of the longitudinal nature of the main study, participation was restricted to freshmen (17.3%), sophomores (36.7%), and juniors (46.0%). Other eligibility criteria included being at least 18 years old, owning a mobile phone with a service contract and text messaging, and drinking at least twice a week over the past month. A total of 88% of the final sample engaged in heavy episodic drinking (drank 4 or more drinks at a sitting for women; 5 or more for men) at least once in the past week, and 74% exceeded National Institute on Alcohol Abuse and Alcoholism (NIAAA) recommendations for weekly drinking (reported drinking 8 or more drinks for women and 15 or more drinks for men in a typical week).

Procedures. Undergraduates ages 18–24 (N = 8,923) were randomly selected from the university registrar's enrollment list and were sent email invitations to participate in an online screening survey to determine eligibility for a longitudinal study of alcohol use and drinking consequences. A total of 3,210 students completed the 15-minute screening survey, for which they were compensated $10. Participants who met eligibility criteria were invited to complete an online 30- to 45-minute baseline survey (compensation $30) that included additional measures of alcohol expectancies, alcohol problems, and drinking consequences. Of the 539 invited, 516 students completed the baseline interview, and 352 enrolled in the longitudinal study.

During the longitudinal study, participants used their mobile phones to complete computerized telephone interviews, conducted with an interactive voice response system. Over the course of a year (four academic quarters, including the summer quarter), participants completed three interviews (morning between 9 A.M. and noon; afternoon 3 P.M.–6 P.M.; evening 9 P.M.–midnight) each day for two randomly assigned weeks in each quarter. Each interview took less than 10 minutes to complete. The analysis described in this article used data from the afternoon interview in which participants answered the 15 expectancy questions (see *Study 2: Measures,* below). Participants were paid $2 for each complete interview, plus a bonus of $16 if they completed 36 of the 42 possible interviews for each 2-week period.

All procedures were approved by the university institutional review board, and a federal Certificate of Confidentiality was obtained from the National Institutes of Health. There were no adverse events reported.

Study 2: Measures

General alcohol outcome expectancies and evaluations were assessed at baseline using the Comprehensive Effects of Alcohol questionnaire (CEOA: Fromme et al., 1993), a 38-item measure of common outcome expectancies assessing the perceived likelihood of various positive and negative effects (4-point scale: 1 = *not at all* to 4 = *very much*) and the perceived evaluation of the effect independent of likelihood (5-point scale: 1 = *bad,* 3 = *neutral,* 5 = *good*). The CEOA has been widely used with college students and has good internal reliability (Cronbach's α ranges from .59 to .89; Fromme and D'Amico, 2000), and test–retest reliability over 3 months (.41–.65, Fromme and D'Amico, 2000). For the purposes of the present study, four subscales were computed: positive expectancies, positive evaluations, negative expectancies, and negative evaluations (α ranged from .85 to .92).

Daily alcohol expectancies and evaluation ratings were measured by asking participants, "If you were to drink tonight, how likely would you be to feel or do the following things? For each item, press a number from 1 to 9, where 1 is very unlikely and 9 is very likely." This instruction was

followed by 15 alcohol-related expectancies. The list was then repeated with the instruction, "If you were to drink tonight and had these experiences, how bad or good would they be? For each item, press a number from 1 to 9, where 1 is extremely bad and 9 is extremely good." For scale items, see Table 1.

Data analyses

The overall goals of the current analyses were to examine item characteristics and scale properties for daily alcohol expectancies and evaluations, taking into account the intensive longitudinal nature of the data. A basic question was whether students would differentiate expectancies (likelihood) and evaluations (good–bad ratings) on a day-to-day basis. Thus, descriptive statistics included not only the mean and standard deviation but also item intraclass correlation coefficients (ICCs), where the ICC is the proportion of the total item variance that is between persons. With intensive longitudinal data, as an ICC trends toward 1, it indicates more traitlike qualities, with little day-to-day variability. Conversely, as it trends toward 0, the ICC indicates more statelike qualities, with little consistency over time within person.

After 2 years of data collection, more than 13,000 records were available for analysis, each representing a single interview containing the expectancy questions. Following descriptive statistics, multilevel exploratory factor analysis (EFA) was used to examine the dimensionality of alcohol expectancies and evaluations at both between-person and within-person levels (Reise et al., 2005). Multilevel EFA is broadly similar to classical EFA but factors the item variance–covariance matrix into between-person and within-person matrices; hence, the dimensionality of the scales can vary across these levels. We considered the use of multilevel confirmatory factor analysis models, in which a specific factor structure is specified in advance (Kline, 2010). Although the current items were selected to broadly tap a positive and negative dimension of alcohol expectancies and evaluations, there has been no prior daily assessment of these constructs. Thus, the multilevel EFA approach was appropriate for these initial analyses. The current analyses used the methods described in Reise et al. (2005) and Hox (2010), which were coded in R (R Development Core Team, 2013), with some comparisons against similar models fit in Mplus (Muthén & Muthén, 1998–2011).

To assess the reliability of the resulting scales, we used generalizability theory (Shavelson & Webb, 1991). Generalizability theory is an extension of reliability from classical test theory that incorporates multiple facets of a study design into the reliability analysis. As an example, internal consistency of a set of items based on Cronbach's alpha must assume that items have no temporal variability, whereas generalizability theory can directly incorporate multiple sources of variance, such as item variability (i.e., internal consis-

tency) and temporal variability. In the present study, the total variance of alcohol expectancy (or evaluations) items can be decomposed into the following variance components:

$$Var(Y) = \sigma_P^2 + \sigma_{D,PD}^2 + \sigma_I^2 + \sigma_{PI}^2 + \sigma_{DI,PDI,e}^2 \quad (1)$$

where P indexes persons, D indexes days, I indexes items, multiple subscripts indicate interactions (e.g., PD represents the interaction of persons and days), and the final term ($\sigma_{DI,PDI,e}^2$) represents the residual variance. Items and persons are crossed factors in that every person completes the exact same set of items, and, as a result, it is possible to estimate the variance of persons by items, which reflects whether certain people systematically differ in certain items. However, days (i.e., repeated measures) represent a nested factor (and are functionally unique within individuals). Because of this, it is not possible to identify the person by day or item by day variances. After estimating the variance terms in Equation 1 via mixed models, they can be used along with the relevant sample sizes to estimate generalizability coefficients (GCs) for both person means and daily means:

$$GC_{Person} = \frac{\sigma_P^2}{\sigma_P^2 + \frac{\sigma_{D,PD}^2}{n_D} + \frac{\sigma_{PI}^2}{n_I} + \frac{\sigma_{DI,PDI,e}^2}{n_D * n_I}} \quad (2)$$

$$GC_{Day} = \frac{\sigma_P^2}{\sigma_P^2 + \sigma_{D,PD}^2 + \frac{\sigma_{PI}^2}{n_I} + \frac{\sigma_{DI,PDI,e}^2}{n_D * n_I}} \quad (3)$$

Similar to Cronbach's alpha, GCs are measures of true variability due to persons (i.e., σ_P^2) as a proportion of total variance. Equations 2 and 3 differ in that in the former we are taking an average GC over days, and thus we divide the days variance term by the number of days.

A final set of analyses examined the convergent and discriminant validity of the new measures relative to well-established cross-sectional measures of expectancies and evaluations. These analyses used a multivariate mixed model (Baldwin et al., 2014). This model included all four newly created daily measures (i.e., positive and negative subscales of expectancies and evaluations) as outcomes simultaneously, predicted from baseline cross-sectional measurements of each construct. All analyses were done in R v3.0.1 (R Development Core Team, 2013) and made use of the lme4 (Bates et al., 2013) and MCMCglmm (Hadfield, 2010) packages for mixed models.

Results

Descriptive statistics

Table 2 contains descriptive statistics for alcohol expectancies and evaluations, and Figure 1 displays medians and

TABLE 2. Descriptive statistics and intraclass correlation coefficients (ICCs) for expectancy and evaluation items

Variable	M (SD)	ICC	[95% CI]
Expectancies			
Relaxed	6.3 (1.83)	.36	[.32, .41]
Sociable	6.6 (1.82)	.31	[.27, .36]
Better mood	6.1 (1.82)	.32	[.28, .36]
Buzz	6.5 (2.17)	.30	[.26, .34]
Energetic	5.7 (1.92)	.36	[.32, .41]
Express feelings	5.4 (1.95)	.39	[.34, .43]
Desire for sex	5.3 (2.18)	.42	[.37, .47]
Unable to study	4.8 (2.68)	.31	[.27, .36]
Hangover	3.6 (2.07)	.36	[.32, .40]
Aggressive	2.4 (1.57)	.36	[.32, .41]
Nauseated or vomit	2.6 (1.62)	.32	[.27, .36]
Hurt/injure self	2.5 (1.56)	.39	[.34, .43]
Unable to remember	3.0 (1.89)	.36	[.31, .41]
Rude/obnoxious	3.0 (1.76)	.40	[.35, .44]
Do something embarrassing	3.4 (1.85)	.35	[.31, .40]
Evaluations			
Relaxed	7.0 (1.60)	.36	[.30, .40]
Sociable	6.6 (1.67)	.33	[.28, .37]
Better mood	6.8 (1.73)	.30	[.26, .34]
Buzz	6.3 (1.88)	.35	[.31, .40]
Energetic	6.2 (1.68)	.29	[.25, .33]
Express feelings	5.2 (1.83)	.39	[.34, .43]
Desire for sex	4.9 (1.95)	.42	[.38, .48]
Unable to study	3.6 (1.88)	.29	[.24, .34]
Hangover	2.5 (1.56)	.27	[.23, .31]
Aggressive	2.4 (1.57)	.39	[.34, .44]
Nauseated or vomit	2.2 (1.43)	.27	[.23, .32]
Hurt/injure self	2.4 (1.49)	.33	[.28, .38]
Unable to remember	2.7 (1.65)	.40	[.35, .45]
Rude/obnoxious	2.6 (1.48)	.38	[.33, .43]
Do something embarrassing	2.8 (1.49)	.36	[.32, .41]

Notes: M = mean; *SD* = standard deviation; CI = confidence interval. Expectancy scores ranged from 1 (*very unlikely*) to 9 (*very likely*), with higher scores representing greater perceived likelihood of occurrence. Evaluation scores ranged from 1 (*extremely bad*) to 9 (*extremely good*), with higher scores representing more positive evaluation.

interquartile ranges. As seen in Figure 1, items that have a general positive connotation were rated highly, indicating both high expectancy and high evaluations, whereas the converse was true for items with general negative connotations. A few items were rated moderately (e.g., not being able to study, desire for sex, and ability to express myself), perhaps indicating more variability across individuals as to whether these items are uniformly positive or negative. Although the mean and median of items differed along a positive to negative dimension, there was wide variability in all items, and all item ranges for both measures encompassed the full scale (i.e., from 1 to 9). Figure 1 also makes clear that there was a very strong correlation between expectancies and evaluations.

Table 2 also includes ICCs and 95% confidence intervals for ICCs for all items. The majority of ICCs were between .30 and .40, indicating that 30% to 40% of total item variability can be attributed to between-person effects (as opposed to variability across days). This in itself is an important finding with the present items, as it was not clear whether repeated daily asking of alcohol expectancies would reveal sufficient within-person variability. These results

suggest that 60% to 70% of the observed item variability is within person. The ICCs also suggest that it is reasonable to pursue multilevel EFA, as there is clear evidence for item variability both within and between persons.

Multilevel exploratory factor analysis

Multilevel EFA was applied to all items, separately for expectancies and evaluations. Models specified one to four factors at both within- and between-person levels, and promax rotation was used before interpreting the factor solutions. Decisions about how many factors to retain were guided by both theory and empirical considerations. From a theory standpoint, items were chosen to broadly reflect two classes of expectancies and evaluations, based on positive and negative connotations. Empirically, we examined scree plots of eigenvalues and factors that contained at least two item loadings at or above .40 (Gorsuch, 1983). Using these criteria, two factors were retained at both within- and between-person levels for both expectancies and evaluations. Factor loadings are presented in Table 3.

The majority of items loaded cleanly on one of two factors across level and measure, as seen in Table 3. There was a clear negative factor, including hangover, becoming aggressive, vomiting, injuring oneself, forgetting, being rude, and embarrassing oneself. Likewise, there was a clear positive factor, including relaxed, sociable, good mood, getting a buzz, energetic, and expressing oneself. Two items were equivocal in their loadings: not being able to study and desire for sex. Thus, for the following analyses, these last two items were excluded, and positive and negative expectancy and evaluations scales included six and seven items, respectively.

Generalizability coefficients

To estimate the GCs of the scales suggested by the multilevel EFA, we first estimated the necessary variance components in Equations 1–3. Linear mixed models were used to estimate these variances after stacking all item responses into a single column with indicator variables that identified which values of the outcome belonged to which items. This is a common technique for fitting multivariate multilevel models (see, e.g., Baldwin et al., 2014). Models were fit separately for positive and negative items of expectancies and evaluations (i.e., four models in total). Applying Equations 2 and 3 to the resulting variance components yielded the GCs found in Table 4. As seen there, on the whole, the positive and negative subscales had excellent reliabilities at both the within- and between-person levels. All but the within-person positive expectancy had GCs of .84 or higher.

The final set of analyses examined convergent and discriminant validity using baseline cross-sectional measures of positive and negative alcohol expectancies and
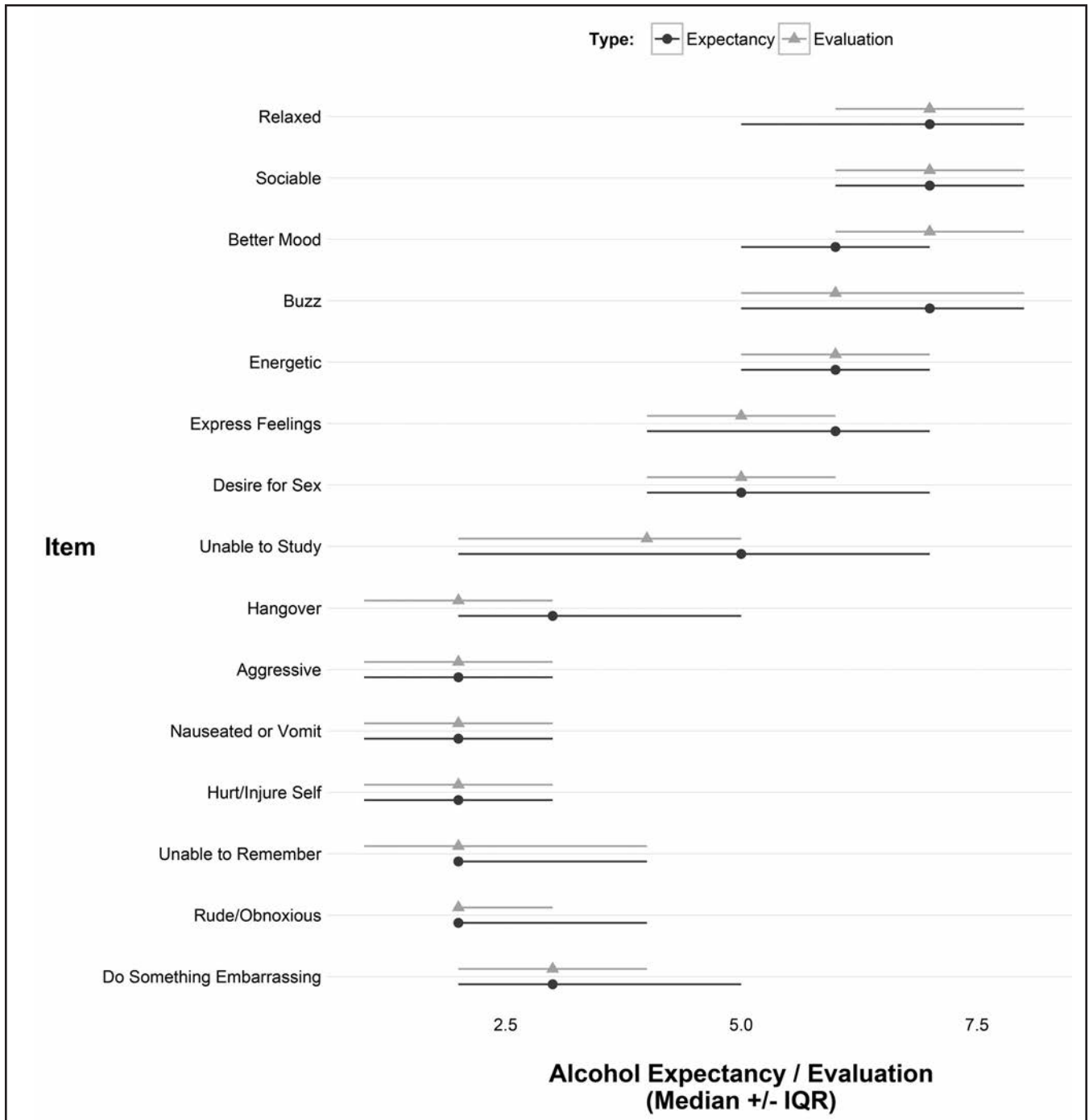
FIGURE 1.   Median and interquartile range (IQR) of expectancy and evaluation items

evaluations (measured with the CEOA). These analyses are presented in Table 5. The results of the multivariate mixed model analyses revealed strong convergent validity, wherein each baseline measure significantly predicted its associated daily measure. In addition, there was evidence of discriminant validity in that cross-valence associations (e.g., positive expectancies at baseline associated with

negative daily expectancies) were either not significant or significantly negatively associated.

Discussion

The current research was designed to develop a psychometrically sound and comprehensive measure of alcohol

TABLE 3.   Factor loadings from multilevel exploratory factor analysis for expectancy and evaluation items

Variable	Between		Within	
	Factor 1	Factor 2	Factor 1	Factor 2
Expectancies				
Relaxed	-.24	.84	.58	-.16
Sociable	-.10	.94	.73	-.08
Better mood	-.14	.96	.71	-.13
Buzz	.12	.67	.58	.16
Energetic	.05	.82	.61	.05
Express feelings	.06	.77	.55	.10
Desire for sex	.26	.50	.46	.14
Unable to study	.39	.37	.26	.28
Hangover	.70	.08	.10	.55
Aggressive	.86	-.16	-.12	.53
Nauseated or vomit	.84	-.14	-.14	.67
Hurt/injure self	.83	-.06	-.08	.61
Unable to remember	.78	-.06	.06	.60
Rude/obnoxious	.93	-.06	-.01	.55
Do something embarrassing	.84	.07	.09	.54
Evaluations				
Relaxed	-.26	.80	-.12	.49
Sociable	-.09	.91	-.07	.59
Better mood	-.15	.92	-.11	.57
Buzz	.09	.74	.04	.54
Energetic	-.02	.90	-.01	.51
Express feelings	.17	.67	.07	.42
Desire for sex	.32	.46	.12	.32
Unable to study	.61	.13	.35	.15
Hangover	.83	-.12	.49	-.02
Aggressive	.88	-.09	.52	-.09
Nauseated or vomit	.87	-.20	.64	-.10
Hurt/injure self	.85	-.05	.58	-.04
Unable to remember	.79	.02	.50	.10
Rude/obnoxious	.93	-.05	.55	-.04
Do something embarrassing	.83	.12	.51	.01

TABLE 4.   Generalizability coefficients for multilevel expectancy and evaluation subscales

	Positive	Negative
Expectancy		
Between	.91	.92
Within	.79	.87
Evaluations		
Between	.91	.95
Within	.84	.91

commonly used comprehensive expectancy measure for college student drinking.

Within- and between-person comparison in expectancies and evaluations

One important implication of the present research was documenting significant within-person variability in alcohol expectancies and evaluations of the effects of alcohol. Results from Study 2 indicate that the selected alcohol expectancies and evaluations items had significant within-person variability (60%–70% at the item level), suggesting that individuals do have daily fluctuations in expectancies and in how they evaluate those expected effects. Whereas the majority of prior work with alcohol expectancies has treated expectancies as stable person characteristics, the present results clearly support within-person variation in alcohol expectancies. This finding has important implications for understanding maintenance of high-risk alcohol use or identifying patterns of naturalistic changes in use. Future research could examine predictors of alcohol expectancies and evaluations, such as current or expected situational context and characteristics, current mood, or planned/intended drinking.

Interestingly, the wide within- and between-person variability along the positive–negative continuum of individual evaluation items suggests that (a) some individuals rate expected effects as more positive or negative in general (i.e., between-person variability; for example, some individuals may rate vomiting as positive on average or relaxed as negative on average) and (b) regardless of general evaluations of expectancies, individuals rate expected effects as more

expectancies that would be appropriate for daily assessments. Our ultimate goal with this measure is to examine a daily process design of alcohol use, alcohol expectancies, and alcohol-related consequences. Results from this study indicate that we have developed a successful 13-item scale with good psychometric properties. Consistent with theory and hypotheses, results from our multilevel EFA indicated that the items loaded onto two factors (positive and negative expectancies/evaluations) and showed excellent reliability at the within- and between-person levels. Last, the measure evidenced good convergent and discriminant validity with a

TABLE 5.   Regression coefficients and 95% CI from multivariate mixed model of daily expectancy and evaluation subscales predicted from baseline CEOA subscales

Variable	Positive		Negative	
	B	[95% CI]	B	[95% CI]
Daily expectancies				
CEOA—Positive, expectancy	0.47	[0.18, 0.73]	-0.93	[-1.24, -0.60]
CEOA—Negative, expectancy	-0.16	[-0.45, 0.15]	0.99	[0.64, 1.33]
Daily evaluations				
CEOA—Positive, evaluation	0.17	[0.03, 0.30]	-0.42	[-0.71, -0.13]
CEOA—Negative, evaluation	-0.09	[-0.24, 0.07]	0.34	[0.08, 0.66]

Notes: CI = confidence interval; CEOA = Comprehensive Effects of Alcohol scale (Fromme et al., 1993).

positive or negative on different days (i.e., significant within-person variability). Future research could also explore antecedents or predictors of these daily fluctuations in evaluations.

Theoretical and clinical implications

Reliable and valid measurement of daily alcohol expectancies has clear theoretical and methodological implications for the substance use disorders field, supporting prospective associations between alcohol expectancies and alcohol use to inform alcohol expectancy theory, as well as learning theory more generally. The ability to assess alcohol expectancies and the evaluations of those expected effects repeatedly over time and in conjunction with resulting alcohol use or nonuse can highlight potential feedback loops or cyclical patterns in use, particularly if one is also considering actual experienced consequences and how they influence future expectancies. Using measurement burst designs, or repeated multiple bursts of measurements over time (Nesselroade, 1991; Sliwinski, 2008), researchers can examine intraindividual change in the within-person relationships between expectancies and alcohol use and whether this may vary as a function of experience with alcohol.

The present research demonstrates that alcohol expectancies are dynamic rather than static. A variety of current interventions targeting college student drinking includes expectancy challenge components alone or in combination with other intervention techniques, as either didactic content presented via video or in-person discussion of alcohol's physiological versus psychological effects (e.g., Corbin et al., 2001; Larimer et al., 2001; Marlatt et al., 1998), or in vivo (experiential) expectancy challenge involving actual alcohol/placebo administration to participants (e.g., Darkes & Goldman, 1993; Musher-Eizenman & Kulick, 2003; Wiers & Kummeling, 2004; Wiers et al., 2005; Wood et al., 2007). Although promising, results of research on the efficacy of alcohol expectancy challenge as a standalone intervention have been mixed. There has been greater support for in vivo interventions over purely didactic ones (see Larimer & Cronce, 2007, for a review). But even across in vivo intervention studies, there have been inconsistent results with respect to the impact of expectancy challenge on indices of alcohol use and consequences and the extent to which changes in expectancies are observable and mediate drinking outcomes. In part, this may be related to the measurement and targeting of expectancies as trait rather than state variables in these interventions and the failure to account for drinking experience and contextual influences on expectancies and their relationship to future drinking.

Limitations

The present results should be viewed in light of certain limitations. First, the sample consisted of only college students (freshmen through juniors) who indicated drinking at least twice per week. Drinking information collected at baseline indicated that this was a high-risk college student sample, as 88% reported heavy drinking in the last week and 74% exceeded NIAAA-recommended guidelines in a typical week. Thus, the results may not generalize to very inexperienced or infrequent drinkers. However, even though our eligibility criterion was drinking twice per week in the past month, examination of drinking patterns in our sample during the course of this expectancy validation shows variability in drinking patterns, including light to heavy patterns of use.

The items in this daily measure do not reflect more serious consequences (e.g., being arrested, alcohol poisoning, driving under the influence). We carefully considered including these items and decided that, although these types of experiences probably have an impact on naturalistic changes in drinking, the infrequency of these consequences makes them less useful in a scale that needed to be brief and relevant for most students on a daily basis. That is, asking if participants think it is likely they will be arrested as a result of drinking this evening probably would not have yielded any variation in responses. The daily interviews also used the same 15 items as a measure of consequences from the previous day's drinking. In addition, they included an open-ended question about whether participants experienced consequences of their drinking other than these 15 consequences. Preliminary coding of those responses did not yield any additional consequences not represented in the 15 items. Furthermore, the measure used in this study evidenced good concurrent and discriminant validity with a widely used alcohol expectancy measure for young adult college student alcohol expectancies.

Conclusions

Alcohol expectancies can vary both across people and within the individual drinker. Instruments used to measure expectancies on a daily basis must be able to capture both within-and between-person variability while various technical and usability issues specific to daily administration are considered. We developed a 13-item scale with good psychometric properties that establish its appropriateness for use in daily process studies of alcohol use.

References

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior.* Englewood Cliffs, NJ: Prentice-Hall.

Armeli, S., Carney, M. A., Tennen, H., Affleck, G., & O'Neil, T. P. (2000). Stress and alcohol use: A daily process examination of the stressor-vulnerability model. *Journal of Personality and Social Psychology, 78,* 979–994.

Armeli, S., Dehart, T., Tennen, H., Todd, M., & Affleck, G. (2007). Daily interpersonal stress and the stressor–vulnerability model of alcohol use. *Journal of Social and Clinical Psychology, 26,* 896–921.

Armeli, S., Mohr, C., Todd, M., Maltby, N., Tennen, H., Carney, M. A., & Affleck, G. (2005). Daily evaluation of anticipated outcomes from alcohol use among college students. *Journal of Social and Clinical Psychology, 24,* 767–792.

Baldwin, S. A., Imel, Z. E., Braithwaite, S. R., & Atkins, D. C. (2014). Analyzing multiple outcomes in clinical research using multivariate multilevel models. *Journal of Consulting and Clinical Psychology, 82,* 920–930.

Bates, D., Maechler, M., & Bolker, B. (2013). *lme4: Linear mixed-effects models using Eigen and S4. R package version 0.999999-2.* Retrieved from http://CRAN.R-project.org/package=lme4

Beatty, P. C., & Willis, G. B. (2007). Research synthesis: The practice of cognitive interviewing. *Public Opinion Quarterly, 71,* 287–311.

Brown, N. R. (1995). Estimation strategies and the judgment of event frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1539–1553.

Brown, S. A., Christiansen, B. A., & Goldman, M. S. (1987). The Alcohol Expectancy Questionnaire: An instrument for the assessment of adolescent and adult alcohol expectancies. *Journal of Studies on Alcohol, 48,* 483–491.

Brown, S. A., Goldman, M. S., Inn, A., & Anderson, L. R. (1980). Expectations of reinforcement from alcohol: Their domain and relation to drinking patterns. *Journal of Consulting and Clinical Psychology, 48,* 419–426.

Burton, S., & Blair, E. (1991). Task conditions, response formulation processes, and response accuracy for behavioral frequency questions in surveys. *Public Opinion Quarterly, 55,* 50–79.

Butler, A. B., Dodge, K. D., & Faurote, E. J. (2010). College student employment and drinking: A daily study of work stressors, alcohol expectancies, and alcohol consumption. *Journal of Occupational Health Psychology, 15,* 291–303.

Corbin, W. R., Bernat, J. A., Calhoun, K. S., McNair, L. D., & Seals, K. L. (2001). The role of alcohol expectancies and alcohol consumption among sexually victimized and nonvictimized college women. *Journal of Interpersonal Violence, 16,* 297–311.

Darkes, J., & Goldman, M. S. (1993). Expectancy challenge and drinking reduction: Experimental evidence for a mediational process. *Journal of Consulting and Clinical Psychology, 61,* 344–353.

Fromme, K., & D'Amico, E. J. (2000). Measuring adolescent alcohol outcome expectancies. *Psychology of Addictive Behaviors, 14,* 206–212.

Fromme, K., Stroot, E., & Kaplan, D. (1993). Comprehensive Effects of Alcohol: Development and psychometric assessment of a new expectancy questionnaire. *Psychological Assessment, 5,* 19–26.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

Hadfield, J. D. (2010). *MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R Package.* Retrieved from cran.r-project.org/web/packages/MCMCglmm/vignettes/Overview.pdf

Ham, L. S., Zamboanga, B. L., Bridges, A. J., Casner, H. G., & Bacon, A. K. (2013). Alcohol expectancies and alcohol use frequency: Does drinking context matter? *Cognitive Therapy and Research, 37,* 620–632.

Helzer, J. E., Badger, G. J., Searles, J. S., Rose, G. L., & Mongeon, J. A. (2006). Stress and alcohol consumption in heavily drinking men: 2 years of daily data using interactive voice response. *Alcoholism: Clinical and Experimental Research, 30,* 802–811.

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). New York, NY: Routledge.

Johnston, L. D., O'Malley, P. M., Bachman, J. G., & Schulenberg, J. E. (2013). *Monitoring the Future national survey results on drug use, 1975–2012: Volume 2, College students and adults ages 19–50.* Ann Arbor, MI: Institute for Social Research, The University of Michigan.

Jones, B. T., Corbin, W., & Fromme, K. (2001). A review of expectancy theory and alcohol consumption. *Addiction, 96,* 57–72.

Jones, B. T., & McMahon, J. (2003). Negative Alcohol Expectancy Questionnaire. In J. P. Allen & V. B. Wilson (Eds.), *Assessing alcohol prob-lems: A guide for clinicians and researchers* (2nd ed., pp. 469–472). Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism. Retrieved from http://pubs.niaaa.nih.gov/publications/AssessingAlcohol/index.pdf

Kline, R. B. (2010). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford Press.

Krank, M., Wall, A. M., Stewart, S. H., Wiers, R. W., & Goldman, M. S. (2005). Context effects on alcohol cognitions. *Alcoholism: Clinical and Experimental Research, 29,* 196–206.

Larimer, M. E., Turner, A. P., Anderson, B. K., Fader, J. S., Kilmer, J. R., Palmer, R. S., & Cronce, J. M. (2001). Evaluating a brief alcohol intervention with fraternities. *Journal of Studies on Alcohol, 62,* 370–380.

Larimer, M. E., & Cronce, J. M. (2007). Identification, prevention, and treatment revisited: Individual-focused college drinking prevention strategies 1999-2006. *Addictive Behaviors, 32,* 2439–2468.

Leigh, B. C., & Stacy, A. W. (1993). Alcohol outcome expectancies: Scale construction and predictive utility in higher order confirmatory models. *Psychological Assessment, 5,* 216–229.

Leigh, B. C., & Stacy, A. W. (1994). Self-generated alcohol expectancies in four samples of drinkers. *Addiction Research & Theory, 1,* 335–348.

MacLatchy-Gaudet, H. A., & Stewart, S. H. (2001). The context-specific positive alcohol outcome expectancies of university women. *Addictive Behaviors, 26,* 31–49.

Mallett, K. A., Bachrach, R. L., & Turrisi, R. (2008). Are all negative consequences truly negative? Assessing variations among college students' perceptions of alcohol related consequences. *Addictive Behaviors, 33,* 1375–1381.

Marlatt, G. A., Baer, J. S., Kivlahan, D. R., Dimeff, L. A., Larimer, M. E., Quigley, L. A., . . . Williams, E. (1998). Screening and brief intervention for high-risk college student drinkers: Results from a 2-year follow-up assessment. *Journal of Consulting and Clinical Psychology, 66,* 604–615.

Mohr, C. D., Armeli, S., Tennen, H., Temple, M., Todd, M., Clark, J., & Carney, M. A. (2005). Moving beyond the keg party: A daily process study of college student drinking motivations. *Psychology of Addictive Behaviors, 19,* 392–403.

Musher-Eizenman, D. R., & Kulick, A. D. (2003). An alcohol expectancy-challenge prevention program for at-risk college women. *Psychology of Addictive Behaviors, 17,* 163–166.

Muthén, L. K., & Muthén, B. O. (1998–2011). *Mplus user's guide* (6th ed.). Los Angeles, CA: Authors.

Nesselroade, J. R. (1991). The warp and woof of the developmental fabric. In R. Downs, L. Liben, & D. S. Palermo (Eds.), *Visions of aesthetics, the environment, and development: The legacy of Joachim F. Wohwill* (pp. 213–240). Hillsdale, NJ: Lawrence Erlbaum Associates.

Perkins, H. W. (2002). Surveying the damage: A review of research on consequences of alcohol misuse in college populations. *Journal of Studies on Alcohol, Supplement 14,* 91–100.

R Development Core Team. (2013). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Reise, S. P., Ventura, J., Nuechterlein, K. H., & Kim, K. H. (2005). An illustration of multilevel factor analysis. *Journal of Personality Assessment, 84,* 126–136.

Shavelson, R. J., & Webb, N. M. (1991). *Generalizability theory: A primer.* Newbury Park, CA: Sage.

Sliwinski, M. J. (2008). Measurement-burst designs for social health research. *Social and Personality Psychology Compass, 2,* 245–261.

Wall, A. M., McKee, S. A., & Hinson, R. E. (2000). Assessing variation in alcohol outcome expectancies across environmental context: An examination of the situational-specificity hypothesis. *Psychology of Addictive Behaviors, 14,* 367–375.

Wall, A.-M., McKee, S. A., Hinson, R. E., & Goldstein, A. (2001). Examining alcohol outcome expectancies in laboratory and naturalistic bar set-

tings: A within-subject experimental analysis. *Psychology of Addictive Behaviors, 15,* 219–226.

Wall, A. M., Thrussell, C., & Lalonde, R. N. (2003). Do alcohol expectancies become intoxicated outcomes? A test of social-learning theory in a naturalistic bar setting. *Addictive Behaviors, 28,* 1271–1283.

Wechsler, H., Dowdall, G. W., Maenner, G., Gledhill-Hoyt, J., & Lee, H. (1998). Changes in binge drinking and related problems among American college students between 1993 and 1997. Results of the Harvard School of Public Health College Alcohol Study. *Journal of American College Health, 47,* 57–68.

Wiers, R. W., & Kummeling, R. H. (2004). An experimental test of an alcohol expectancy challenge in mixed gender groups of young heavy drinkers. *Addictive Behaviors, 29,* 215–220.

Wiers, R. W., van de Luitgaarden, J., van den Wildenberg, E., & Smulders, F. T. Y. (2005). Challenging implicit and explicit alcohol-related cognitions in young heavy drinkers. *Addiction, 100,* 806–819.

Willis, G. B. (2005). *Cognitive interviewing: A tool for improving questionnaire design.* Thousand Oaks, CA: Sage.

Wood, M. D., Capone, C., Laforge, R., Erickson, D. J., & Brand, N. H. (2007). Brief motivational intervention and alcohol expectancy challenge with heavy drinking college students: A randomized factorial study. *Addictive Behaviors, 32,* 2509–2528.