2017

Psychometric analysis and validity of the daily alcohol-related consequences and evaluations measure for young adults

Scott A. Baldwin
Brigham Young University - Provo

Christine M. Lee

Jessica M. Cronce

Anne M. Fairlie

David C. Atkins

See next page for additional authors

HHS Public Access

Author manuscript

Psychol Assess. Author manuscript; available in PMC 2018 March 01.

Published in final edited form as:

Psychol Assess. 2017 March ; 29(3): 253–263. doi:10.1037/pas0000320.

Psychometric Analysis and Validity of the Daily Alcohol-related Consequences and Evaluations Measure for Young Adults

Christine M. Lee[1], **Jessica M. Cronce**[1], **Scott A. Baldwin**[2], **Anne M. Fairlie**[1], **David C. Atkins**[1], **Megan E. Patrick**[3], **Lindsey Zimmerman**[4], **Mary E. Larimer**[1], and **Barbara C. Leigh**[5]

[1]Department of Psychiatry and Behavioral Sciences, University of Washington, Box 354944, Seattle, Washington, 98195.

[2]Department of Psychology, Brigham Young University, 268 TLRB, Provo, UT, 84602.

[3]Institute for Social Research, University of Michigan, 426 Thompson St., Ann Arbor, MI 48106-1248.

[4]National Center for PTSD Dissemination and Training Division, VA Palo Alto Health Care System, Menlo Park, CA 94025.

[5]Alcohol and Drug Abuse Institute, University of Washington, Box 354805, Seattle, Washington, 98195.

Abstract

College students experience a variety of effects resulting from alcohol use and evaluate their experiences on a continuum from negative to positive. Using daily reports collected via cell phone, we examined the psychometric properties of alcohol use consequences and evaluations of alcohol use consequences. Participants were 349 undergraduate students (mean age 19.7 [SD=1.26], 53.4% female). Data were analyzed using a multilevel factor analysis framework, incorporating binary items (consequences) and normally distributed items (evaluations). Our model converged on two factors - positive and negative - with similar loadings between- and within-persons. Intraclass correlation coefficients (ICC) for positive consequences and their evaluations ranged from .30-.40, whereas values for negative consequences were more variable. ICCs for negative evaluations were higher, suggesting evaluations were more trait-like compared to experience of consequences which may be context dependent. Generalizability coefficients (GC) on the whole were good to excellent, suggesting highly reliable scales at both person mean and daily mean levels. However, likely due to binary scale and infrequency, the GC for negative consequences at the daily level was somewhat low. Convergent validity was demonstrated by (1) positive associations between baseline RAPI and AUDIT scores with latent factors for daily positive and negative consequences, and (2) positive associations between daily drinking and daily consequences and evaluations of consequences. Overall, this measure demonstrated good psychometric properties for use in studies examining daily and lagged relationships between alcohol use and related consequences.

Correspondence concerning this article should be addressed to Christine M. Lee, Department of Psychiatry and Behavioral Sciences, University of Washington, Box 354694, Seattle, WA 98195. leecm@u.washington.edu.

Jessica M. Cronce is now in the Department of Counseling Psychology and Human Services in the College of Education at the University of Oregon.

Keywords

Alcohol-related consequences; daily questionnaire; psychometrics

Alcohol use and its associated consequences have been widely studied in young adult and college student populations. The term *consequence* has a definitively negative connotation, most often referring to the damage caused by alcohol misuse (e.g., Hingson, Zha, & Weitzman, 2009; Perkins, 2002), which impacts the individual drinker (e.g., hangovers, blackouts) as well as others (e.g., perpetration of physical assault and sexual violence). For some (but not all) drinkers, identifying previously experienced consequences motivates behavior change (see Cronce & Larimer, 2011, for a review of brief motivational interventions). This is consistent with an operant learning theory perspective, where consequences are assumed to be *punishers*, and their occurrence is expected to decrease undesirable drinking behavior (Bouton, 2000). The theories of reasoned action (Fishbein, 1979) and planned behavior (Ajzen, 1991) suggest that attitudes about behaviors, including evaluations of whether a consequence is good or bad, are considered when forming behavioral intentions and predicting behavior.

Evaluations of alcohol consequences are, however, not uniform (Patrick & Maggs, 2011). For example, vomiting and blackouts are routinely included in measures of negative consequences and used to predict decreases in drinking, but close to 25% and 50% of students, respectively, consider these consequences to be positive or neutral (Mallett, Bachrach & Turrisi, 2008). Understanding how and when certain consequences act as *reinforcers* (i.e., positive consequences that tend to increase behavior) may be just as, if not more, important in predicting and intervening in drinking behavior (Patrick & Maggs, 2008). Of course, evaluations of whether or not a given consequence was experienced as positive, neutral, or negative may vary over time as circumstances change. For example, having a hangover on a day that you can sleep in is likely less negative than having a hangover the morning you have to take a test.

It is important to consider how students evaluate potential future consequences as well as experienced past consequences. Research supports the relation between college students' evaluations of potential *future* consequences (i.e., how positive [good] or negative [bad] a consequence would be, regardless of their intent to drink) and alcohol use and problems (Gaher & Simons, 2007; Neighbors, Walker, & Larimer, 2003; Patrick & Maggs, 2011). Experiencing more positive consequences is associated with evaluating future positive consequences as more positive; experiencing more negative consequences is associated with perceiving future consequences as less negative (Logan, Henry, Vaughn, Luk, & King, 2012). Other research has focused on students' evaluations of *experienced* consequences, namely their evaluations of how positive or negative a consequence was based on their actual experience of it (Merrill, Read, & Barnett, 2013; Merrill, Read, & Colder, 2013; White & Ray, 2014). However, prior studies have chiefly relied on retrospective assessment methods that are subject to recall bias. Moreover, they fail to account for individual variation in evaluations of consequences over time (see Merrill, Read & Barnett, 2013), and often fail to

obtain reports on all of the consequences experienced (e.g., they focus only on negative consequences).

In fact, prior research considering the experience of both positive and negative consequences simultaneously is scant (e.g., Park, 2004; Park & Grant, 2005; Park & Levenson, 2002; Patrick & Maggs, 2008). Examining positive consequences in addition to negative consequences seems critical, as students report experiencing positive consequences significantly more frequently than negative consequences across drinking occasions (Park, 2004; Park & Grant, 2005; Park & Levenson, 2002; Patrick & Maggs, 2008), and alcohol consumption is related to the extent to which students' most positive consequence is considered positive, but not to the extent that their most negative consequence is negative (Park, 2004). Unfortunately, extant studies examining positive consequences have used measures that classified consequences as positive a priori, consistent with extant negative consequence measures (e.g., Hurlbut & Sher, 1992; Kahler, Strong, & Read, 2005; White & Labouvie, 1989).

Building off our own and others' prior work, we developed a program of research to examine a daily process model of alcohol use, alcohol expectancies, experienced consequences, and evaluations of experienced consequences (Corbin, Morean, & Benedict, 2008; Lee et al., 2015). The daily process model accounts for different types of associations among the key constructs, including alcohol expectancies predicting later day alcohol use and experienced consequences, as well as feed-forward processes which may help describe the maintenance or change in future alcohol use (e.g., consequences experienced today may predict next-day expectancies and alcohol use). Daily process models using daily reports of behaviors and psychological constructs, like evaluations of experienced consequences, are important for several reasons. First, daily process models can distinguish between-person effects from within-person effects. Between-person effects refer to person-level characteristics that may partially explain variation in an outcome, such as gender partially accounting for variation in the number of drinks consumed on a given day. Within-person effects refer to characteristics (e.g., situational or psychological) that may differ from occasion to occasion for a given individual, and these characteristics can partially account for day-to-day changes in an outcome. For instance, daily fluctuations in negative mood may partially explain day-to-day changes in alcohol consumption for a given individual over time. Second, daily process models allow for the examination of the temporal ordering of psychological constructs and/or behaviors. Daily reports of evaluations of negative consequences may predict reductions in drinking the following day when the consequences are perceived as especially negative (see Merrill, Read, & Barnett, 2013). Third, unlike retrospective reports, daily reports are less adversely affected by the inability to remember events or feelings. Daily reports also minimize bias that may be introduced by the passage of time, as in the case of re-evaluating an experience after a few weeks have passed.

To date, there has not been an established psychometrically-sound daily measure of experienced consequences. Using a multilevel factor analysis framework, incorporating up to 56 days of data within person across 349 individuals, the present study establishes the psychometric properties of a 13-item daily positive and negative consequences measure, which includes evaluations of experienced consequences, at both the between- and within-

person levels, while taking into account the key features of this type of data including: (a) repeated measures nested within individuals, (b) different item scales (i.e., binary items for consequences and continuous scales for evaluations), and (c) evaluation scores being contingent upon whether an individual consequence was experienced.

To psychometrically evaluate the proposed measure, we proceeded in four steps. First, we evaluated the intraclass correlations of each consequence and evaluation to determine how much items varied between- and within-persons. Large between-person variance indicates that consequences and evaluations are fairly constant across time for a given person. Conversely, large within-person variance indicates that consequences and evaluations fluctuate notably within-persons, over time. Second, we conducted separate exploratory multilevel factor analyses to examine between- and within-person factors for consequences and evaluations. We hypothesized two factors (i.e., positive and negative factors) both between- and within-person for consequences and evaluations. Third, based on results of the previous step, we used a multilevel confirmatory factor analysis, which included the latent factors for both consequences and evaluations simultaneously at both between- and within-person levels of the data. This model was necessary in it allowed us to account for the fact that participants only rated a consequence if they reported experiencing the consequence. Thus, we obtained unbiased estimates of factor loadings, factor variances, and factor correlations. With respect to between-factor correlations, we hypothesized that positive consequences would be positively associated with positive evaluations and negative consequences at both levels. We anticipated that negative consequences would be negatively associated with negative evaluations (i.e., greater endorsement of negative consequences is associated with less favorable evaluations), at both between- and within-person levels. Finally, convergent validity was be assessed by extending the prior confirmatory factor analysis model to include baseline predictors (i.e., negative consequences as measured by the Rutgers Alcohol Problem Index [White & Labouvie, 1989] and the Alcohol Use Disorders Identification Test [Babor, Higgins-Biddle, Saunders, & Monteiro, 2001]) of the latent variable constructs for the daily-level consequences and evaluations. We hypothesized that the proposed daily measure of alcohol-related consequences and evaluations of consequences would demonstrate good convergent validity with relevant baseline measures of alcohol-related consequences and daily measures of alcohol use.

Methods

Participants

Participants for the present analyses were part of a larger study examining a daily process model of alcohol use, alcohol expectancies and consequences and included 349 undergraduate students (mean age 19.7 [SD=1.26], 53.4% female). Due to the longitudinal nature of the larger research study, enrollment was only open to students of freshman, sophomore, and junior standing. Most participants (74.2%) were Caucasian, with the remainder Asian American (8.5%), multiracial (11.1%), or other (6.2%).

Procedures

Undergraduate freshman, sophomore, and junior students between the ages of 18-24 (N=8,923) at a large public university were randomly selected from the University Registrar's enrollment list and invited to participate in a larger longitudinal daily study of alcohol use and related consequences. Students were invited via email and mailed letter to complete a short confidential screening survey to assess study eligibility, which included owning a mobile phone with a service contract and text messaging, being at least 18 years old, and drinking at least twice a week over the past month. Those who met eligibility criteria were then invited to participate in an online baseline survey assessing demographics, alcohol use, and other psychosocial measures. Of the 3,210 students who completed the screening survey, 539 met criteria and were invited to the baseline survey. Of those, 516 completed the baseline survey, and 352 came to an in-person training session at the study offices, which included consent procedures, as well as an overview of the study procedures and training in the data collection method, and were enrolled in the longitudinal study (i.e., consented and completed at least 1 daily interview). Of the 164 participants who completed baseline and did not enroll in the study, 58 individuals declined to come in for a training session, 1 completed the training but did not start the daily interviews, and 105 either never attended their scheduled training session or did not schedule a training session. The main reasons participants cited for missing a scheduled training session and/or not rescheduling a session was that they were too busy or that the available times did not work for them. No significant differences were found between those eligible participants who enrolled vs. didn't enroll in the study based on age, $t(364.42) = -1.48$, $p = .14$, gender, χ^2 (1, N = 516) = 1.19, $p = .28$, total drinks per week, $t(489) = -0.66$, $p = .51$, AUDIT sum scores, $t(500) = 0.47$, $p = .64$, and negative consequences, $t(260.65) = -0.33$, $p = .74$, at baseline. Participants were compensated $10 for completing the screening survey and $30 for completing the baseline assessment.

Participants used their mobile phones to complete daily telephone interviews via an Interactive Voice Response system. Participants completed three interviews a day for four bursts of 2-week periods over the course of one academic year. Daily interviews included a morning interview (9am-noon), afternoon interview (3pm-6pm) and evening interview (9pm-midnight). Each interview took less than 10 minutes to complete and participants were compensated $2 for each complete interview, plus a bonus of $16 if they completed 36 of the 42 possible interviews for each 2-week period. Of the 352 participants who came in for a training session, three participants did not report drinking during the daily portion of the study, resulting in a sample of 349 for the present analyses.

All procedures were approved by the University IRB and a federal Certificate of Confidentiality was obtained from the National Institutes of Health.

Measures

Daily alcohol-related consequences and evaluations (DACE) for young adults

—Item development and selection of consequences occurred in conjunction with the development of a daily alcohol expectancy measure (Lee et al., 2015), based on review of the alcohol expectancies and alcohol effects literatures, expert review of items, and cognitive

interviews with college students (see Lee et al., 2015, for specific details of the selection of initial items). We conducted cognitive interviews with 14 college student drinkers, exploring whether the proposed consequence items were representative of the alcohol effects that college students would experience and what important effects might have been missed.

The purpose of the broader study was to examine a daily process model of alcohol expectancies, alcohol use, and related consequences and we opted to measure the same consequence effects as expectancies, due to the fact that the measure of consequences needed to be brief, assess both positive and negative consequences and evaluations, and evidence sufficient between- and within-person variability (with a varying level of endorsement throughout the entire study). For example, severe consequences (e.g., alcohol poisoning, getting arrested due to drinking and driving) may occur and have influence on future drinking and alcohol expectancies; however, the base rates for these consequences are comparatively low and would contribute zero or very little variability between- and within-person, and would therefore not be useful toward the development of a daily measure of consequences. Thus, for the purpose of the present analyses, we used 15 alcohol-related consequences and evaluations.

In each morning interview, participants who reported drinking the previous day were asked, “Did any of the following things happen to you as a result of your drinking yesterday?” Participants could respond yes or no to 15 different alcohol-related consequence items (e.g., I felt relaxed, I became aggressive). For the consequences endorsed, the list was then repeated with the question, “How bad or good was that, from 1 to 9, where 1 is extremely bad and 9 is extremely good?” (see Table 1 for scale items). An open-ended question regarding whether there were any additional consequences that occurred that day as a result of their drinking was also included, as a way to catch more serious consequences. Preliminary coding of those consequences did not yield any consequences that were not represented in the 15 items.

Daily alcohol use—In the morning interview, participants reported on their alcohol use on the previous day. Participants were asked “Did you drink any alcohol yesterday, from the time you got up to the time you went to sleep?” For those who reported drinking, they were then asked “How many drinks did you have in total yesterday?” Participants reviewed standard drink definitions during the initial training session.

Baseline alcohol consequences—Participants were asked to indicate how many times in the last three months they had experienced each of 23 alcohol-related consequences, as measured by the Rutgers Alcohol Problem Index (White & Labouvie, 1989). Items were dichotomized to reflect whether the participant experienced the item in the past three months (1) or not (0). Items were summed together to create the number of different alcohol-related consequences the participant had experienced in the last three months at baseline.

Baseline high-risk alcohol use—The Alcohol Use Disorders Identification Test (AUDIT, Babor et al., 2001) was administered at baseline to provide a measure of harmful or hazardous alcohol use. The AUDIT consists of 10 items covering consumption, drinking behavior/dependence, and alcohol-related problems, and the items were summed to create a

total score. The AUDIT has been found to have reasonable psychometric properties among samples of college students for use in determining high-risk drinking among college students (Kokotailo et al., 2004).

Data Analyses

The present analyses focused on the psychometric and substantive evaluation of two types of daily reports: a) alcohol-related consequences, and b) evaluations of the consequences, when experienced. Thus, the data analyses had to incorporate several aspects of the resulting data, including (a) repeated measures nested within individuals, (b) differing item scales (i.e., binary items for consequences and continuous scales for evaluations), and (c) the absence of an evaluation score if the consequence did not occur. On the latter point, the evaluation data can be thought of as having a structural missing data pattern, or as subject to a selection process (Kim & Muthén, 2009). The evaluation occurs only after a certain threshold has been passed, namely that the related consequence occurred.

Data were analyzed using a multilevel factor analysis framework, allowing for between- and within-person factors and associations of latent variables. In addition, given the mixture of binary and continuous items, the final model is a type of multimodal multilevel model, in which the likelihood incorporates both binary and normally distributed data. Finally, the joint model functions like a selection model, in which a consequence must first be experienced prior to an evaluation being observed, and the factor structure and loadings of evaluation items are conditional on the consequences. Due to its complexity, the resulting model was fit using a fully Bayesian analysis, including minimally informative priors and Markov chain Monte Carlo (MCMC) estimation (Song & Lee, 2012).

The reliability of the resulting scales was examined with generalizability coefficients (GC; Shavelson & Webb, 1991), which are an extension of classic internal reliability to research designs with multiple sources of error (e.g., longitudinal data). Generalized linear mixed models were used to estimate the following variance components:

$$Var\,(Y) = \sigma^2_P + \sigma^2_{D,PD} + \sigma^2_I + \sigma^2_{PI} + \sigma^2_{DI,PDI,e} \quad (1)$$

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 represents the residual variance.[1] Items and persons are crossed factors as each person completes the same set of items; thus, 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. Using the variance terms from Equation 1, GCs for both person means and daily means can be estimated via:

[1]Consequence items were analyzed using a logistic mixed model, which does not include an error term. For these scales, the variance of the logistic distribution ($\frac{\pi^2}{3}$) was used for the residual error.

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

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

Similar to Cronbach's alpha, GCs are measures of true variability due to persons as a proportion of total variance, where the additional terms in the denominators of Equations 2 and 3 represent error variance. Equations 2 and 3 differ in that in the former we are taking an average GC over days, and thus divide the days-variance term by the number of days. Analyses were conducted using Mplus v7.11 (Muthén & Muthén, 2012) and R v3.0.1 (R Development Core Team, 2013).

Results

Descriptive Statistics

Due to larger study aims of examining daily process models of alcohol use, expectancies and consequences, our goal was to recruit frequently drinking college students who would report experiencing a variety of alcohol-related consequences on any given occasion. Of the 352 participants who completed the baseline survey, completed the in-person training session, and began the longitudinal daily diary portion of the study, 88% 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 at baseline, and 74% exceeded NIAAA recommendations for weekly drinking (reported drinking 8 or more drinks for women and 15 or more drinks for men in a typical week). On average, participants drank 18.82 (*SD* = 11.60) total drinks per week in the past three months and experienced 11.71 (*SD* = 8.79) negative consequences in the last three months at baseline. Further, the average score on the AUDIT was 13.72 (*SD* = 5.48).

Descriptive statistics for consequences and evaluations can be found in Table 1. As seen in the table, items with generally positive connotations (e.g., being more social) were endorsed more frequently and rated more positively, whereas items with generally negative connotations (e.g., vomiting) were less frequent and rated less positively. Figure 1 contains a dot plot of intraclass correlation coefficients (ICCs) for each item for both consequences and consequence evaluations. The ICC is the proportion of the total item variance that is attributed to the variance between people, and hence values close to one reflect strong, trait-like qualities, whereas values close to zero indicate highly variable responses within individuals. As seen in the figure, positive items are relatively consistent with ICCs between .30 and .40 across both consequences and their evaluations. Negative items' ICCs are notably more variable. Interestingly, for several items the consequence ICC is quite low, whereas the corresponding evaluation ICC is quite high (e.g., injuring oneself, being aggressive). This would suggest that the consequence itself is more context dependent (i.e.,

varies notably within people over time), but a person's evaluation of the consequence is more stable.

Exploratory multilevel factor analyses were used separately on consequences and evaluations.[2] Based on fit indices and interpretability, results suggested a two factor model at both between- and within-person levels. Fit indices for the two-factor model were: RMSEA = 0.02, CFI = 0.95, TLI = 0.93, and SRMR = 0.06. By way of comparison, the fit indices for the one-factor model at both levels were: RMSEA = 0.04, CFI = 0.78, TLI = 0.75, and SRMR = 0.13. Most items cleanly loaded on one of two factors, largely conforming to positive vs. negative alcohol-related consequences. Two exceptions to this pattern were the items “Being unable to study” and “Having more desire for sex.” Item loadings for these items were generally lower, and in one instance similar across factors. The loadings for these two items were similar to earlier analyses focused on alcohol expectancies and expectancy evaluations (Lee et al., 2015), and they were not included in the following analyses.

The multimodal, multilevel factor model described earlier was fit to the remaining items and is shown in Figure 2. Based on the preceding exploratory factor analyses, the joint model included positive and negative latent factors for both consequences and consequence evaluations at both between-person and within-person levels of the data. Maximum likelihood methods typically run until a model has “converged”—that is, until there is minimal change in the estimates. Bayesian methods do not have such a criterion, rather an explicit number of iterations must be specified and the resulting estimates must be evaluated for convergence. Current analyses used the Gelman-Rubin diagnostic and traceplots (Gelman & Hill, 2006) to assess convergence of the posterior distributions of parameter estimates, and these diagnostics were consistent with convergence. Factor loadings from the Bayesian model (i.e., mean of the posterior distribution) and 95% credible intervals (CI) are shown in Figure 3 and listed in Tables 2 and 3. All factor loadings are significantly different than zero, though the positive items are estimated much more precisely relative to the negative items. The width of the CI is directly related to the amount of information in the data, where positive consequences were reported much more often, and hence positive consequences evaluations were observed more frequently. Negative consequences were reported far less often, and their evaluations were observed less frequently. Nonetheless, the point estimates show that items reliably load on their corresponding factors and that in general, items have similar loadings between- and within-persons.

The multimodal, multilevel factor model also provides variance-covariance matrices of the latent factors for both between- and within-levels of the model. A dotplot of factor correlations and CI is presented in Figure 4, which shows an interesting pattern of associations among consequences and their evaluations. Not surprisingly, positive consequences and their evaluations show a strong, positive association. Similarly, positive and negative consequences are positively correlated. Between persons, there is no association between evaluations of positive consequences and evaluations of negative

[2]The exploratory factor analysis for consequence evaluations had difficulties converging due to sparse data. Results from the joint model described later also strongly suggest two factors.

consequences. However, within persons (i.e., in day-to-day fluctuations) there is a small, positive correlation between these two different types of evaluations. Finally, within persons, the number of negative consequences and their evaluation had a moderately strong negative correlation. Between-persons negative consequence endorsement and their evaluation were not associated. However, the difference in these two correlations is moderately large and bounded away from zero (r_{diff} = .62, 95% CI = [.27, .99]).

GCs were used to estimate the internal consistency of scales for both person means and daily means and are found in Table 4. GCs on the whole were good to excellent, suggesting highly reliable scales at both person mean and daily mean levels. The one exception to this was that the GC for negative consequences at the daily level was somewhat low. This reflects in part the scaling of consequences (i.e., binary) as well as their somewhat infrequent nature.

Convergent Validity

We examined the convergent validity of the daily-level consequences and evaluation scales in two ways. First, we extended the confirmatory factor analysis model to include baseline predictors of the latent variable constructs for the daily-level consequences and evaluations. Specifically, baseline RAPI and AUDIT scores were included as predictors of the between-persons consequences and evaluation latent variables (see Figure 2). Baseline RAPI was significantly and positively related to positive consequences ($\hat{\beta}=0.02$, 95% CI = 0.01, 0.04) and negative consequences ($\hat{\beta}=0.04$, 95% CI = 0.03, 0.06). These coefficients as well as the other validity coefficients can be interpreted as follows. For the positive consequences latent variable, which is standardized with a mean of 0 and standard deviation of 1, a one-unit increase in baseline RAPI is associated with a 0.02 standard deviation increase in positive consequences. For the negative consequences latent variable, a one-unit increase in baseline RAPI is associated with a 0.04 standard deviation increase in negative consequences. One could also interpret the relationships in terms of standard deviation change in the RAPI. For example, a 1 standard deviation increase in baseline RAPI (*SD* = 8.8) is associated with a 0.4 (8.8 × 0.04) standard deviation increase in negative consequences.

Baseline RAPI was not significantly related to positive evaluations ($\hat{\beta}=0.01$, 95% CI = −0.001, 0.03) or negative evaluations ($\hat{\beta}=-0.02$, 95% CI = −0.05, 0.01). The relationships between the RAPI and positive and negative evaluations were in the expected direction, but not significant. This was likely due to the fact that there was less data for evaluations than the consequences, because participants only evaluated the consequences when the consequences occurred. Baseline AUDIT was significantly and positively related to positive consequences ($\hat{\beta}=0.03$, 95% CI = 0.01, 0.05) and negative consequences ($\hat{\beta}=0.07$, 95% CI =0.04, 0.10). Baseline AUDIT was not significantly related to positive evaluations ($\hat{\beta}=-0.001$, 95% CI =−0.03, 0.02) or negative evaluations ($\hat{\beta}=-0.01$, 95% CI =−0.06, 0.04).

The second method for evaluating convergent validity was examining whether daily drinking predicted scale scores for consequences and evaluations. The positive and negative consequences scale scores were the sum of positive or negative consequences, respectively,

on a given day. The evaluations of positive and negative consequences were the mean positive or negative evaluations, respectively, on a given day. When predicting daily consequences, we used a multilevel Poisson model with robust standard errors. When predicting daily evaluations, we used a multilevel normal model.

Daily drinking had a positive relationship with positive consequences (incident rate ratio [IRR] = 1.07, $p < 0.01$, 95% CI = 1.06, 1.08) and negative consequences (IRR = 1.23, $p < 0.01$, 95% CI = 1.21, 1.26). Thus, a 1-drink increase in daily drinking is associated with a 7% increase in the number of positive consequences and 23% increase in the number of negative consequences experienced on a given day. Daily drinking also had a positive relationship with positive evaluations ($\hat{\beta}=0.04$, $p < 0.01$, 95% CI = 0.03, 0.05), showing consuming more alcohol was associated with more favorable evaluations of positive consequences. Daily drinking had a negative relationship with negative evaluations ($\hat{\beta}=-0.03$, $p < 0.01$, 95% CI = −0.05, −0.01), showing consuming more alcohol was associated with less favorable evaluations of negative consequences.

Discussion

The current research was designed to develop and examine the psychometric properties of a daily positive and negative alcohol-related consequences measure. Our ultimate goal with this measure is to use it to examine the daily and lagged relationships with alcohol use and alcohol-related consequences. Results from multilevel factor analyses support a 13-item scale with good psychometric properties. Two factors emerged, representing a subscale for positive consequences and a subscale for negative consequences at both between- and within-person level of data. Reliabilities of the two subscales were high, despite the psychometric difficulties associated with binary items that assess infrequent consequences. In addition, there was variance in the experience and evaluation of positive and negative consequences both between-persons on average and within-individuals across days. The findings from the present study support the use of this new measure in research utilizing intensive repeated measures designs.

The present results offer some interesting descriptive findings. Consistent with other cross-sectional research (Park, 2004; Park & Grant, 2005; Park & Levenson, 2002; Patrick & Maggs, 2008), we found that positive consequences were endorsed more frequently and more positively across measurement occasions. With the exception of feeling energetic and expressing feelings more, all the positive items were endorsed 50% or more of the time. Negative consequences were much less frequently endorsed, despite this being a college sample recruited for drinking at least twice per week.

The current results demonstrate that both the likelihood of occurrence and the evaluation of consequences vary both between and within individuals. This finding has implications for alcohol interventions based on operant learning theories (Bouton, 2000; Monti, Kadden, Rohsenow, & Abrams, 2002; Marlatt & Donovan, 2005), which assume behavior is shaped by its consequences. For example, understanding variability in likelihood of different consequences, as well as individual evaluations of those consequences, is important for utilizing this information in motivational interventions (Cronce & Larimer, 2011). In

addition, this information is important to more broadly understand how, when, and under what circumstances consequences impact subsequent drinking. Surprisingly, we found there was less variability in how students evaluated the negative consequences, particularly for hurting or injuring oneself by accident and forgetting what they did while drinking, than there was in how they evaluated the positive consequences. This suggests that evaluations for negative consequences may be more dependent on the person or on prior learning experiences and learning contexts than on the immediate context or extremity of the consequence, while positive evaluations are more situational. If negative consequences are presumed to operate as punishers that would decrease subsequent drinking behavior, it is valuable to understand whether more trait-like negative consequence evaluations are associated with same day or next day drinking. Further research on these daily processes in conjunction with alcohol use measures would shed light on whether experiences of specific consequences are particularly aversive and/or for whom they are more likely to be associated with decreases in alcohol use. Identifying associations between negative consequences and within-person changes in alcohol use behaviors would help to identify opportune times and situations in which to intervene.

The present results should be evaluated in light of study limitations. The sample consisted of college students of freshman, sophomore, and junior standing from one university. Further, the sample inclusion criteria limited the participants to those who drank at least twice per week in the last month and to those owning a cell phone with a text messaging plan. Thus the results may not generalize to older individuals, those not in college, or to those with less experience with alcohol. It should be noted, however, that despite our initial drinking inclusion criteria, examination of drinking patterns over the course of the year indicated light to heavy patterns of use. Additional limitations include that the temporal relationship of the reinforcing consequences (e.g., getting a buzz) and the punishing consequences (e.g., having a hangover) differ, with the rewarding consequences more proximal to the behavior. This presents a challenge for future research. Finally, for the purposes of our larger study, we opted to include a very brief measure of consequences and evaluations with item selection based on alcohol effects that would evidence within-person (i.e., occasion-level variability) as well as between-person variability. The measure presented here does not reflect all positive or negative consequences that college students may experience, and in particular does not include serious alcohol-related negative consequences, which might have immediate and long-lasting changes to one's health, well-being, and future decisions about alcohol use (e.g., alcohol poisoning, alcohol-related traffic accidents, alcohol-related arrests).

We developed a scale to measure the consequences of drinking alcohol with several novel features: the inclusion of both positive and negative consequences and their evaluations, a focus on scale items that describe consequences of discrete drinking episodes, and the tailoring of scale items to be used in daily measures of drinking and its consequences. The scale demonstrated both within- and between-person variability in the consequences reported and the evaluations of those consequences, indicating both trait-like and state-like variability and demonstrating the utility of the scale in examining the relationships between alcohol use and alcohol-related consequences across days.

Acknowledgments

Data collection and manuscript preparation were supported by a grant from the National Institute on Alcohol Abuse and Alcoholism (R01 AA016979). The content of this manuscript is solely the responsibility of the author(s) and does not necessarily represent the official views of the National Institute on Alcohol Abuse and Alcoholism or the National Institutes of Health.

References

Ajzen I. The theory of planned behavior. Organizational Behavior and Human Decision Processes. 1991; 50:179–211.

Babor, TF., Higgins-Biddle, JC., Saunders, JB., Monteiro, MG. AUDIT The alcohol use disorders identification test: guidelines for use in primary care. 2nd Edition.. World Health Organization; Geneva: 2001. WHO/MSD/MSB/01.6a

Bouton ME. A learning theory perspective on lapse, relapse, and the maintenance of behavior change. Health Psychology. 2000; 19(1S):57. Doi; 10.1037//0278-6133.19.1(Suppl.).57. [PubMed: 10709948]

Corbin WR, Morean ME, Benedict D. The Positive Drinking Consequences Questionnaire (PDCQ): validation of a new assessment tool. Psychology of Addictive Behaviors. 2008; 33:54–68.

Cronce JM, Larimer ME. Individual-focused approaches to the prevention of college student drinking. Alcohol Research and Health. 2011; 34:210–221. [PubMed: 22330220]

Fishbein M. A theory of reasoned action: Some applications and implications. Nebraska Symposium on Motivation. 1979; 27:65–116.

Gaher RM, Simons JS. Evaluations and expectancies of alcohol and marijuana problems among college students. Psychology of Addictive Behaviors. 2007; 21:545–554. [PubMed: 18072837]

Gelman, A., Hill, J. Data analysis using regression and multilevel/hierarchical models. Cambridge University Press; 2006.

Hingson RW, Zha W, Weitzman ER. Magnitude of and trends in alcohol-related mortality and morbidity among US college students ages 18-24, 1998-2005. Journal of Studies on Alcohol and Drugs. 2009; 16:12–20.

Hurlbut SC, Sher KJ. Assessing alcohol problems in college students. Journal of American College Health. 1992; 41:49–58. [PubMed: 1460173]

Kahler CW, Strong DR, Read JP. Toward efficient and comprehensive measurement of the alcohol problems continuum in college students: The Brief Young Adult Alcohol Consequences Questionnaire. Alcoholism: Clinical and Experimental Research. 2005; 29:1180–1189.

Kim Y, Muthén B. Two-part factor mixture modeling: Application to an aggressive behavior measurement instrument. Structural Equation Modeling. 2009; 16:602–624. [PubMed: 20717486]

Kokotailo PK, Egan J, Gangnon R, Brown D, Mundt M, Fleming M. Validity of the alcohol use disorders identification test in college students. Alcoholism: Clinical and Experimental Research. 2004; 28:914–920.

Lee CM, Atkins DC, Cronce JM, Walter T, Leigh BC. A daily measure of positive and negative alcohol expectancies and evaluations: Documenting a two-factor structure and within- and between-person variability. Journal of Studies on Alcohol and Drugs. 2015; 76:326–335. [PubMed: 25785808]

Logan DD, Henry T, Vaughn M, Luk JW, King KM. Rose-Colored Beer Goggles: The Relation between Experiencing Alcohol Consequences and Perceived Likelihood and Valence. Psychology of Addictive Behaviors. 2012; 26:311–317. [PubMed: 21639598]

Mallett K, Bachrach R, Turrisi R. Are all negative consequences truly negative? Assessing variations among college students' perceptions of alcohol related consequences. Addictive Behaviors. 2008; 33:1375–1381. [PubMed: 18639384]

Merrill JE, Read JP, Barnett NP. The way one thinks affects the way one drinks: Subjective evaluations of alcohol consequences predict subsequent change in drinking behavior. Psychology of Addictive Behaviors. 2013; 27:42–51. [PubMed: 22985054]

Merrill JE, Read JP, Colder CR. Normative perceptions and past-year consequences as predictors of subjective evaluations and weekly drinking behavior. Addictive Behaviors. 2013; 38:2625–2634. [PubMed: 23899424]

Marlatt, GA., Donovan, DM., editors. Relapse Prevention: Maintenance Strategies in Addictive Behaviors. Guilford Press; New York, NY, US:

Monti, P., Kadden, R., Rohsenow, D., Cooney, N., Abrams, D. Treating Alcohol Dependence: A Coping Skills Training Guide (2Nd Ed.) [e-book]. Guilford Press; New York, NY, US: 2002. Available from: PsycINFO, Ipswich, MA. [January 16, 2015]

Muthén, LK., Muthén, BO. Mplus user's guide (7th ed). Muthén & Muthén; Los Angeles, CA: 2012.

Neighbors C, Walker DD, Larimer ME. Expectancies and evaluations of alcohol effects among college students: Self-determination as a moderator. Journal of Studies on Alcohol and Drugs. 2003; 64:292–300.

Park CL. Positive and negative consequences of alcohol consumption in college students. Addictive Behaviors. 2004; 29:311–321. [PubMed: 14732419]

Park CL, Grant C. Determinants of positive and negative consequences of alcohol consumption in college students: Alcohol use, gender, and psychological characteristics. Addictive Behaviors. 2005; 30:755–765. [PubMed: 15833579]

Park CL, Levenson MR. Drinking to cope among college students: Prevalence, problems and coping processes. Journal of Studies on Alcohol. 2002; 63:486–497. [PubMed: 12160108]

Patrick ME, Maggs JL. College students' evaluations of alcohol consequences as positive and negative. Addictive Behaviors. 2011; 36:1148–1153. [PubMed: 21855224]

Patrick ME, Maggs JL. Short-term changes in plans to drink and importance of positive and negative alcohol consequences: Between- and within-person predictors. Journal of Adolescence. 2008; 31:307–321. [PubMed: 17651796]

Perkins HW. Surveying the damage: A review of research on consequences of alcohol misuse in college populations. Journal of Studies on Alcohol. 2002; (Suppl. 14):91–100.

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria: 2013. URL http://www.R-project.org/

Shavelson, RJ., Webb, NM. Generalizability Theory: A Primer. Sage; Newbury Park, CA: 1991.

Song, X., Lee, S. Basic and advanced Bayesian structural equation modeling with applications in the medical and behavioral sciences. Wiley; West Sussex, UK: 2012.

White H, Labouvie E. Towards the assessment of adolescent problem drinking. Journal of Studies on Alcohol. 1989; 50:30–37. [PubMed: 2927120]

White HR, Ray AE. Differential evaluations of alcohol-related consequences among emerging adults. Prevention Science. 2014; 15:115–124. [PubMed: 23412945]

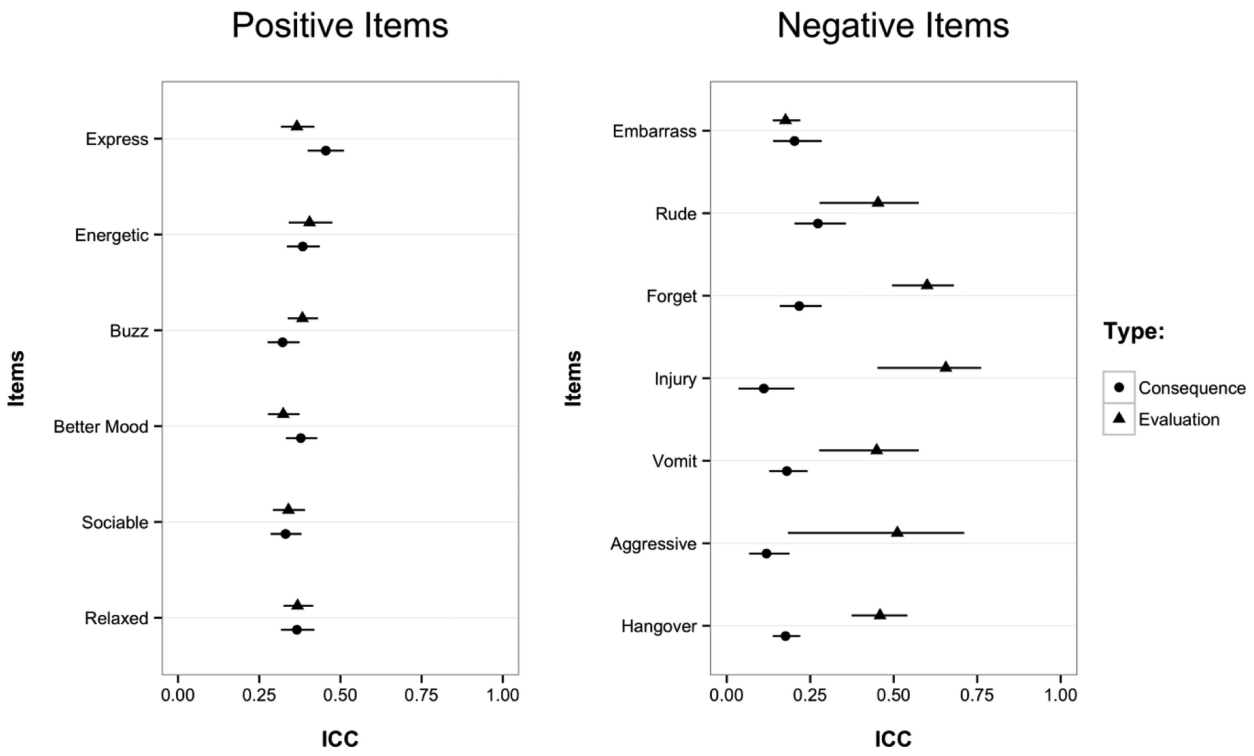

Figure 1.
Intraclass correlations (ICC) of positive and negative consequences and their evaluations.

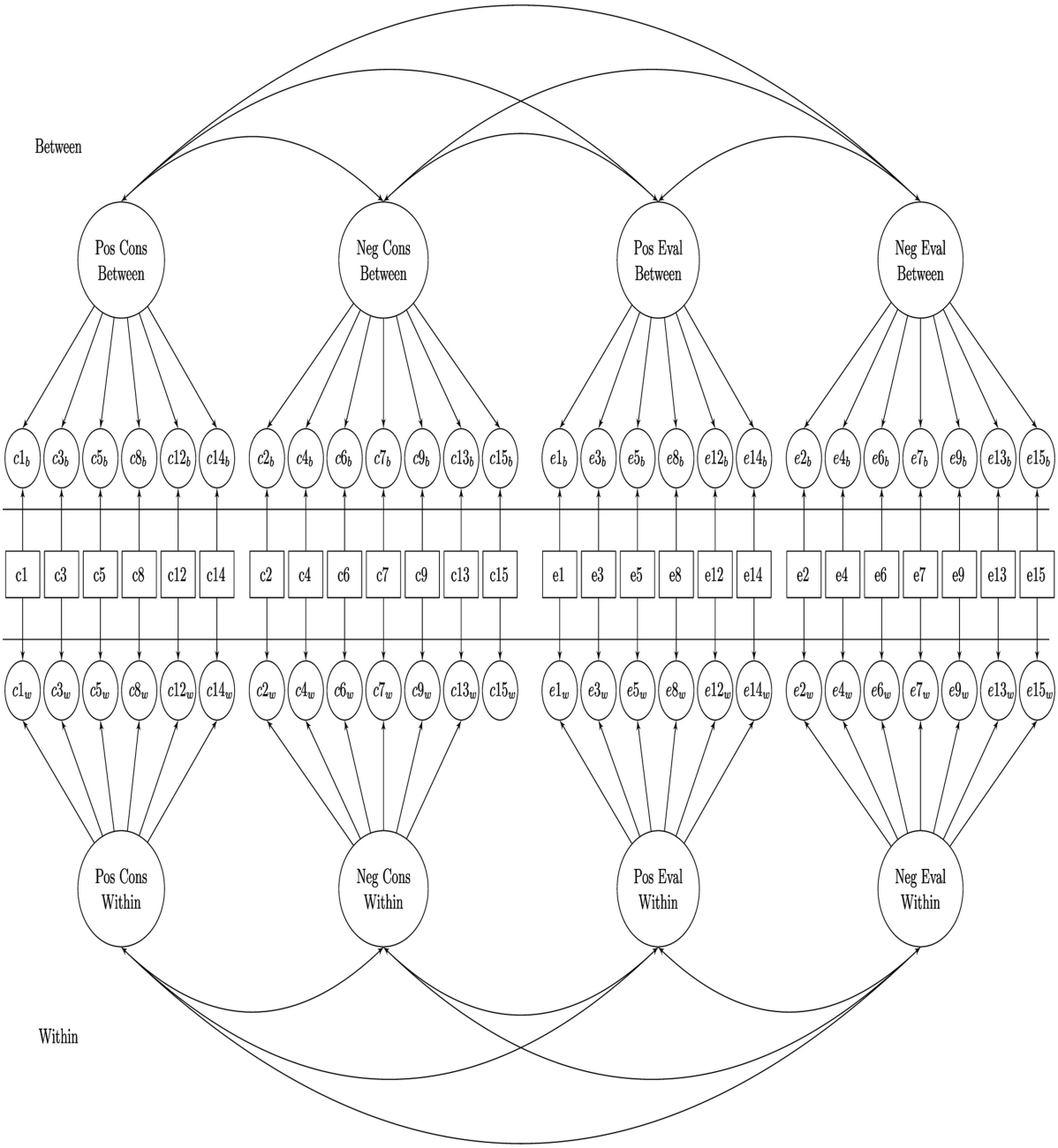

Figure 2.

Path diagram for the multilevel, multimodal confirmatory factor model.

Note. Rectangles represent observed variables and circles/ovals represent latent variables. Labels beginning with a ‘c’ represent consequences and labels beginning with ‘e’ represent evaluations. Numbers within labels represent the items. The items are: 1 “Relaxed”, 2 “Hangover”, 3 “Sociable”, 4 “Aggressive”, 5 “Better Mood”, 6 “Vomit”, 7 “Injury”, 8 “Buzz”, 9 “Forget”, 12 “Energy”, 13 “Rude”, 14 “Express”, and 15 “Embarrass”.

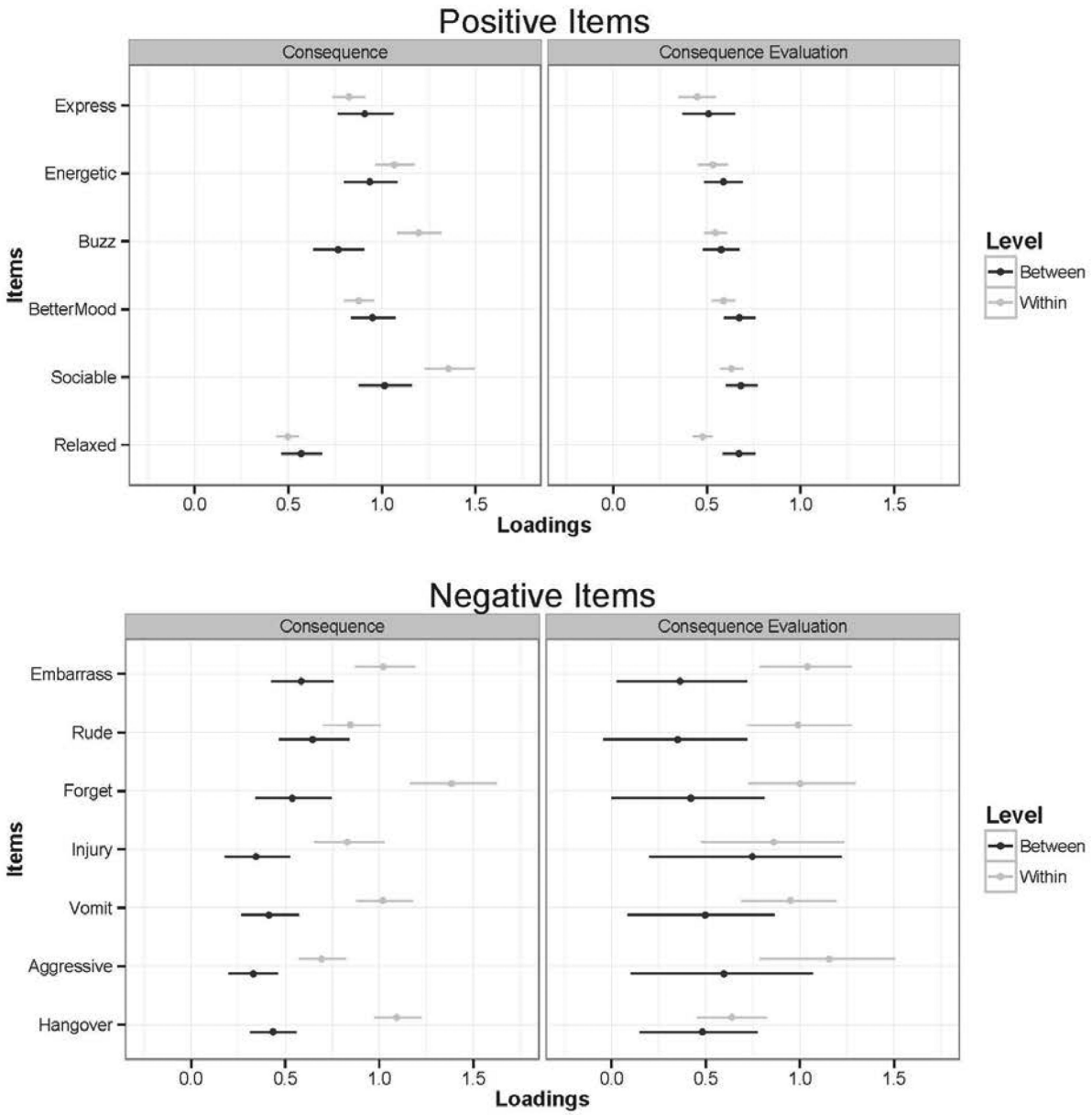

Figure 3.
Factor loadings and credible intervals for positive and negative consequences and their evaluations.

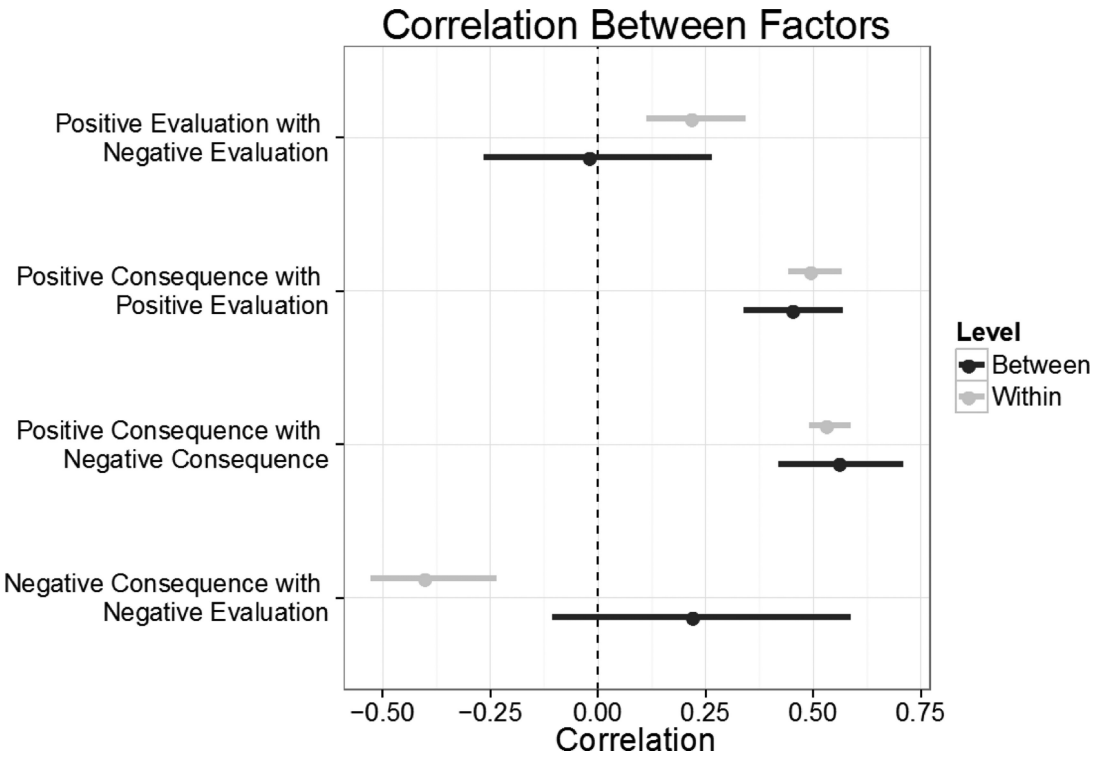

Figure 4.
Factor correlations between positive and negative consequence and their evaluations.

Table 1

Descriptive statistics for consequence items and consequence evaluation items

	Consequences	Evaluations
Stem Questions: **Consequences: "Did any of the following things happen to you as a result of your drinking yesterday?" / Evaluations: "How good or bad was that?"**	**% Endorsed one or more days during daily study**	**Mean (SD)**
Positive Items		
I was able to express my feelings more easily / Expressing my feelings more easily	27	6.0 (1.5)
I felt more energetic / Feeling more energetic	34	6.6 (1.3)
I got a buzz / Getting a buzz	66	6.6 (l.4)
I was in a better mood / Being in a better mood	53	6.9 (1.3)
I was more sociable / Being more sociable	57	6.8 (l.3)
I felt relaxed / Feeling more relaxed	67	6.8 (l.3)
Negative Items		
I did something that embarrassed me / Doing something that embarrassed me	6	3.5 (1.7)
I was rude or obnoxious / Being rude or obnoxious	5	3.7 (1.7)
I couldn't remember what I did while drinking / Forgetting what I did while drinking	7	3.7 (1.8)
I hurt or injured myself by accident / Hurting or injuring myself by accident	2	3.4 (1.8)
I felt nauseated or vomited / Feeling nauseated or vomiting	7	2.9 (1.6)
I became aggressive / Becoming aggressive	4	4.0 (2.3)
I had a hangover / Having a hangover	22	3.6 (1.8)
Dropped Items		
I was unable to study / Being unable to study	26	4.2 (1.8)
I had more desire for sex / Having more desire for sex	25	5.8 (1.8)

Note. N = 349 participants with 5375 observations. Evaluations for all items range from 1 (extremely bad) to 9 (extremely good).

Table 2

Unstandardized Loadings

	Consequence		Consequence Evaluation	
	Loading	**95% CI**	**Loading**	**95% CI**
Between				
Positive Items				
Express	0.91	0.77, 1.06	0.51	0.37, 0.65
Energetic	0.94	0.80, 1.08	0.59	0.49, 0.69
Buzz	0.77	0.64, 0.90	0.58	0.48, 0.67
GoodMood	0.95	0.84, 1.07	0.67	0.59, 0.76
Sociable	1.01	0.88, 1.16	0.68	0.60, 0.77
Relaxed	0.57	0.46, 0.68	0.67	0.59, 0.76
Negative Items				
Embarrass	0.59	0.43, 0.76	0.36	0.03, 0.72
Rude	0.65	0.47, 0.84	0.35	−0.04, 0.72
Forget	0.54	0.34, 0.74	0.42	0.00, 0.81
InjureSelf	0.35	0.18, 0.52	0.75	0.20, 1.22
Vomit	0.41	0.27, 0.57	0.50	0.08, 0.86
Aggressive	0.33	0.20, 0.46	0.60	0.10, 1.07
Hangover	0.44	0.32, 0.56	0.48	0.15, 0.77
Within				
Positive Items				
Express	0.82	0.74, 0.91	0.45	0.35, 0.55
Energetic	1.07	0.97, 1.17	0.53	0.45, 0.61
Buzz	1.20	1.08, 1.32	0.54	0.48, 0.60
GoodMood	0.88	0.80, 0.96	0.59	0.53, 0.65
Sociable	1.36	1.23, 1.49	0.63	0.57, 0.69
Relaxed	0.50	0.44, 0.56	0.48	0.43, 0.53
Negative Items				
Embarrass	1.09	0.97, 1.22	1.04	0.79, 1.27
Rude	0.69	0.57, 0.82	0.99	0.72, 1.28
Forget	1.02	0.88, 1.18	1.00	0.73, 1.30
InjureSelf	0.83	0.65, 1.03	0.86	0.47, 1.24
Vomit	1.38	1.16, 1.62	0.95	0.69, 1.20
Aggressive	0.85	0.70, 1.01	1.16	0.79, 1.51
Hangover	1.09	0.97, 1.19	0.64	0.45, 0.82

Note. N = 349 participants with 5375 observations. The point estimate for the factor loading is the mean of the posterior distribution. 95% CI = 95% credible interval.

Table 3

Standardized Loadings

	Consequence		Consequence Evaluation	
	Loading	**95% CI**	**Loading**	**95% CI**
Between				
Positive Items				
Express	.77	.70, .83	.84	.78, .89
Energetic	.82	.75, .87	.88	.83, .92
Buzz	.72	.64, .79	.88	.83, .92
GoodMood	.89	.84, .93	.71	.62, .78
Sociable	.87	.81, .91	.81	.73, .88
Relaxed	.66	.57, .74	.59	.46, .70
Negative Items				
Embarrass	.76	.63, .86	.46	.04, .79
Rude	.77	.63, .88	.44	−.06, .79
Forget	.58	.41, .73	.42	.00, .74
InjureSelf	.57	.33, .76	.63	.19, .89
Vomit	.59	.42, .73	.53	.10, .84
Aggressive	.61	.41, .77	.55	.10, .85
Hangover	.63	.49, .75	.53	.17, .81
Within				
Positive Items				
Express	.63	.59, .67	.35	.28, .43
Energetic	.73	.69, .76	.46	.39, .51
Buzz	.77	.73, .80	.48	.43, .52
GoodMood	.66	.62, .69	.52	.47, .57
Sociable	.80	.78, .83	.55	.51, .60
Relaxed	.44	.40, .49	.46	.41, .51
Negative Items				
Embarrass	.71	.66, .77	.69	.55, .81
Rude	.64	.57, .71	.64	.49, .79
Forget	.81	.76, .85	.68	.51, .83
InjureSelf	.64	.55, .72	.64	.38, .83
Vomit	.71	.66, .76	.67	.50, .80
Aggressive	.57	.50, .64	.57	.40, .73
Hangover	.74	.70, .77	.41	.30, .53

Note. N = 349 participants with 5375 observations. The point estimate for the factor loading is the mean of the posterior distribution. 95% CI = 95% credible interval. Loadings are standardized within-levels but not across levels. Consequently, direct comparisons are possible within-levels (i.e., comparing between-person loadings to other between-person loadings), but not across loadings (i.e., comparing between-person loadings to within-person loadings). A fully standardized solution is not possible with multimodal, multilevel data. The standardized loadings were computed as: $\sqrt{b^2/\left(b^2\left(\sigma_l^2\right)+\sigma_r^2\right)}$, where $\boldsymbol{b}$ is the unstandardized factor loading, σ_l^2 is the latent variable variance (fixed to 1 for in the models, and σ_r^2 is the residual variance as the between or within levels. σ_r^2 was fixed to 1 for the within-level model for the consequences and it was estimated for all other models.

Table 4

Generalizability coefficients for between-person and within-person scales for consequences and consequence evaluations

	Positive	Negative
Consequences		
Between	.91	.89
Within	.78	.66
Evaluations		
Between	.90	.88
Within	.87	.78

Note. N = 349 participants with 5375 observations.