The Alcohol Sensitivity Questionnaire: Evidence for Construct Validity

Kimberly A. Fleming, Bruce D. Bartholow, Joseph B. Hilgard, Denis M. McCarthy, Susan E. O’Neill, Douglas Steinley, & Kenneth J. Sher

University of Missouri and Midwest Alcoholism Research Center

To appear in Alcoholism: Clinical and Experimental Research. (Note that this is not the version of record, which will be made available from the journal publisher’s website.)

Author Note

Kimberly A. Fleming, Bruce D. Bartholow, Joseph B. Hilgard, Denis M. McCarthy, Susan E. O’Neill, Douglas Steinley, & Kenneth J. Sher, Department of Psychological Sciences and the Midwest Alcoholism Research Center.

This research was supported by grants P60 AA011998, T32 AA013526, F31 AA022551, R01 AA023248 and K05-AA017242 from the National Institute on Alcohol Abuse and Alcoholism.

Correspondence concerning this article should be directed to Bruce D. Bartholow, Department of Psychological Sciences, 210 McAlester Hall, University of Missouri, Columbia, MO 65211. Email: bartholowb@missouri.edu.
Abstract

Background. Variability in sensitivity to the acute effects of alcohol is an important risk factor for the development of alcohol use disorder (AUD). The most commonly used retrospective self-report measure of sensitivity, the Self-Rating of the Effects of Alcohol form (SRE), queries a limited number of alcohol effects and relies on respondents’ ability to recall experiences that might have occurred in the distant past. Here, we investigated the construct validity of an alternative measure that queries a larger number of alcohol effects, the Alcohol Sensitivity Questionnaire (ASQ), and compared it to the SRE in predicting momentary subjective responses to an acute dose of alcohol. Method. Healthy young adults (N = 423) completed the SRE and the ASQ and then were randomly assigned to consume either alcohol or a placebo beverage (between-subjects manipulation). Stimulation and sedation (Biphasic Alcohol Effects Scale) and subjective intoxication were measured multiple times after drinking. Results. Hierarchical linear models showed that the ASQ reliably predicted each of these outcomes following alcohol but not placebo consumption, provided unique prediction beyond that associated with differences in recent alcohol involvement, and was preferred over the SRE (in terms of model fit) in direct model comparisons of stimulation and sedation. Conclusions. The ASQ compared favorably with the better-known SRE in predicting increased stimulation and reduced sedation following an acute alcohol challenge. The ASQ appears to be a valid self-report measure of alcohol sensitivity and therefore holds promise for identifying individuals at-risk for AUD and related problems.

Key words: Alcohol sensitivity; level of response; subjective alcohol effects; alcohol challenge; model comparison
Substantial evidence suggests that risk for alcohol use disorder (AUD) is conferred via sensitivity to the effects of alcohol (e.g., Newlin & Thompson, 1990; Quinn & Fromme, 2011; Schuckit, 1994). Alcohol sensitivity is defined as the amount of alcohol one must consume in order to experience a given effect, or the extent to which a given alcohol dose influences subjective feelings (Pollock, 1992) and physiological (e.g., hormonal, neural) responses (e.g., Schuckit et al., 1987). Since the first demonstration that low sensitivity (LS) at age 20 is associated with substantially greater likelihood of developing an AUD by age 30 (Schuckit, 1994), empirical work on the correlates of alcohol sensitivity has proliferated (for reviews, see Morean & Corbin, 2010; Quinn & Fromme, 2011). Evidence suggests that LS-associated risk is dissociable from other AUD predictors, including alcohol expectancies, externalizing behavior, comorbid psychiatric disorders, and personality (e.g., Trim et al., 2009; Schuckit et al., 2004).

Given these considerations, the ability to easily and reliably measure sensitivity is very important. Ideally, alcohol sensitivity would be assessed through a combination of subjective (e.g., self-reported intoxication) and objective (e.g., standing ataxia; physiological) responses to a laboratory alcohol challenge (see Schuckit, 1994). However, this mode of assessment is cost-prohibitive, and is inappropriate for certain populations who cannot be ethically administered alcohol, such as underage drinkers, individuals with active AUDs, and individuals taking certain medications (Wood & Sher, 2000). Furthermore, laboratory-based assessment is untenable for large-scale epidemiological studies that rely on broadly generalizable and relatively brief instruments.

Self-Rating of the Effects of Alcohol Form

To meet these challenges, Schuckit and colleagues developed the Self-Rating of the Effects of Alcohol (SRE) form (Schuckit et al., 1997). The SRE asks respondents to indicate the
number of drinks required to experience up to four effects from drinking alcohol (recognition of
“any effect;” dizziness or slurred speech; stumbling gait; passing out) during three different time
periods (their first five drinking episodes, period of heaviest drinking in their lives, and most
recent consecutive three month period in which they drank), and to respond only to effects that
were actually experienced in a given timeframe. The SRE has demonstrated good internal
consistency ($\alpha > .90$) and test-retest reliability ($r = .82$) (Schuckit et al., 1997). Concurrent
validity has been established by correlating SRE scores with subjective effects during laboratory
alcohol challenge (Schuckit et al., 1997) and with scores on AUD diagnostic instruments (see
Ray et al., 2011). SRE scores also predict development of AUD and problems prospectively
(e.g., Schuckit et al., 2006, 2007, 2011; Schuckit & Smith, 2001).

Thus, the SRE has been important in advancing understanding of the role of alcohol
sensitivity in the etiology of AUD. Nevertheless, the SRE suffers from some limitations. First,
the scope of effects assessed by the SRE is small and consists primarily of sedation-like
symptoms generally associated with large alcohol doses. Although its brevity likely reduces
subject burden, this factor also likely limits the range of individuals for whom SRE scores fully
reflect drinking experiences. This situation can result in fewer endorsed effects for some
individuals than for others, which can produce skewed estimates of sensitivity level due to an
inherent correlation between the number of effects endorsed and the number of drinks needed to
experience them (Lee et al., 2015). On the other end of the severity spectrum, “feeling any
different” is a relatively vague item that could be open to numerous interpretations, potentially
limiting its utility (see Clark & Watson, 1995).

Another limitation of the SRE is that it requires respondents to recall experiences that
may have occurred many years in the past, or that in any case might be difficult to remember.
Given the problems associated with accurately recalling alcohol use experiences (e.g., Del Boca & Darkes, 2003; Parra et al., 2003), it is likely that the retrospective reports queried by the SRE are less accurate than those related to more proximal experiences.

**Alcohol Sensitivity Questionnaire**

To address these limitations, O’Neill and colleagues (2002) created the Alcohol Sensitivity Questionnaire (ASQ). In creating the ASQ, O’Neill and colleagues aimed to sample a wide range of effects that could be experienced across numerous contexts on both the ascending and descending limbs of the blood alcohol concentration curve. Like the SRE, the ASQ asks respondents to indicate the number of drinks they must consume in order to experience alcohol-related effects. Specifically, the ASQ contains 15 items (see Table 1), of which nine tap effects typically associated with lower doses and stimulation (e.g., feeling more talkative; more flirtatious) and six tap effects typically associated with heavier doses and sedation (e.g., feeling nauseous, passing out). For each item, respondents indicate whether or not they have experienced the effect from drinking alcohol; for each endorsed effect, they estimate the minimum number of drinks they must consume in order to experience the effect (for lower dose/light drinking effects) or the maximum number of drinks they could consume without experiencing the effect (for larger dose/heavy drinking effects). These differing referents are designed to provide estimates of limits on sensitivity across the spectrum of common alcohol effects.

High ASQ scores (indicating LS) are associated with heavy alcohol use (Bartholow et al., 2003; 2007; 2010) and alcohol-related negative consequences (Bartholow et al., 2010; Fleming & Bartholow, 2014). Other evidence linking ASQ scores with AUD risk has come from research showing that ASQ scores uniquely predict heavy drinking prospectively, beyond the influence of
baseline alcohol involvement (Bartholow et al., 2007). Moreover, high ASQ scores are
associated with enhanced brain responses to alcohol-related images (Bartholow et al., 2007,
2010; Shin et al., 2010) but not to other appetitive stimuli (Bartholow et al., 2010).
To date, no study has examined whether ASQ scores predict responses to alcohol
challenge, an essential component of the measure’s construct validity. Further, the ASQ has not
been compared directly with the SRE to determine its performance relative to this better-known
measure. This was the purpose of the current study, for which three primary hypotheses were
advanced. First, consistent with the modified differentiator model (King et al., 2011), we
predicted that higher ASQ scores would be associated with increased feelings of stimulation
following alcohol consumption. Given that the ASQ contains items specifically tapping
stimulation-related effects, which are largely unassessed by the SRE, we predicted that a model
based on ASQ lighter-drinking items would be preferred (in terms of model fit) over an SRE-
based model of these effects. Second, we predicted that higher ASQ scores would predict
reduced feelings of sedation post-consumption. Further, because the ASQ assesses a broader
range of sedating effects, we predicted that a model based on ASQ heavy-drinking items would
be preferred over an SRE-based model in predicting sedation. Finally, we predicted that higher
scores on the ASQ and SRE would be associated with decreased feelings of subjective
intoxication.

Method

Participants
Four hundred fifty-eight adults aged 21-34 (M age = 23.31; 49% female, 88% Caucasian)
were recruited from the Columbia, MO community for a study examining effects of alcohol on
cognition. Study announcements were placed in mass email blasts and in online classifieds.
Interested individuals were instructed to contact the lab. Potential participants were interviewed via telephone; individuals reporting conditions contraindicating participation in an alcohol challenge (abstention; history of alcohol or drug abuse treatment or other serious mental or physical illness; deliberate attempts to cut down on drinking; prescription medication other than oral contraception; pregnancy) or that would impede completion of laboratory tasks (color-blindness; a primary language other than English) were excluded from the sample. In addition, to ensure that the alcohol dose received in the study would be within participants’ normal range of experience, naive drinkers (< 2 drinks per week on average) and very heavy drinkers (≥ 25 drinks per week on average) were excluded from the sample. Eligible individuals were scheduled for the first of two laboratory sessions. Participants received $35 for the baseline session and $14/hour for participation in the second (beverage administration) session.

Self-report Measures

Means and SDs of the measures described in this section, as a function of beverage group assignment, are reported in Table 2.

**Alcohol Sensitivity Questionnaire (ASQ).** The first nine of the ASQ’s 15 items query effects of alcohol often associated with lighter drinking. For each of these items, respondents are asked to indicate whether they have ever experienced the effect as a result of drinking alcohol, and if so, to estimate the minimum number of drinks they need to consume in order to feel the effect. The remaining items, assessing effects most associated with heavier drinking, are structured similarly, except that respondents are asked to estimate the maximum number of drinks they can consume without experiencing the effect.2

Confirmatory factor analysis (CFA) was used to compare a one-factor model of the ASQ to a two-factor model (9 lighter-drinking items, 6 heavier-drinking items). The two-factor model
represented a significant improvement in fit ($\chi^2$ [Difftest] = 881.51, $df = 1$, $p < .001$). Initially, fit for the two-factor model was fair ($\chi^2 = 374.22$, $df = 89$; CFI = .88, RMSEA = .09). Modification indices suggested a significant correlation between the error terms of two items on the lighter-drinking factor, “Sleepy” and “Sluggish.” Given the conceptual similarity between these items, a correlation was specified between their error terms, resulting in a final version of the two-factor model that fit adequately, ($\chi^2 = 271.3$, $df = 88$; CFI = .92, RMSEA = .07). Internal consistency in the current sample was excellent for both factors (ASQ-Heavy $\alpha = .95$; ASQ-Light $\alpha = .89$).

Factors scores were used for primary data analyses.

**Self-Rating of the Effects of Alcohol (SRE) Form.** Respondents indicate the number of standard drinks required to experience up to four different effects (recognition of “any effect;” dizziness or slurred speech; stumbling gait; passing out) over three different time periods (their first five drinking episodes; the period of heaviest drinking in their lives; the most recent consecutive three months in which they drank). In order to approximate the time frame queried by the ASQ, only responses from the most recent consecutive three months of drinking (SRE 3-mo.) were used in the current analyses.

CFA was used to estimate a single factor model of the SRE. Fit indices were mixed, with the CFI indicating a good fit, but with a high RMSEA value ($\chi^2 = 40.37$, $df = 2$; CFI = .95, RMSEA = .21). Internal consistency for the SRE 3-mo. items was good ($\alpha = .83$). As with the ASQ, factor scores were used for primary data analyses.

**Alcohol use and consequences.** Participants reported their average number of drinking occasions per week and average number of drinks consumed per occasion in the past 3 months (scored on a per week basis), using items adapted from the NIAAA Task Force recommendations (NIAAA, 2003). An alcohol quantity-frequency variable (AlcQF) was created...
by multiplying the number of typical weekly drinking occasions by number of drinks typically
consumed per occasion. Participants indicated their experience of various alcohol-related
negative consequences using the 24-item Young Adult Alcohol Problems Screening Test (see
Hurlbut & Sher, 1992). Nine of these items specifically query features of AUD (e.g.,
withdrawal; continued use despite problems). Participants indicated whether they had
experienced each consequence “Never;” “Yes, but not in the past year;” “In the past year but not
the past 3 months;” “Yes, in the past 3 months: once; twice; 3 times; 4 or more times” (scored 0,
.3, .5, 1, 2, 3, and 5, respectively). An overall “negative consequences” score was calculated as
the sum of responses to all 24 items (α = .86); a separate “AUD” score was calculated as the sum
of responses to the nine dependence-related items (α = .75).

Subjective Effects of Alcohol

Stimulation and sedation. The BAES (Martin et al., 1993) is a self-report measure of
stimulant and sedative effects of drinking alcohol. Respondents use a 10-point scale to rate the
extent to which they are currently experiencing seven states associated with sedation (e.g., down,
sluggish) and seven states associated with stimulation (e.g., up, excited). As is customary, BAES
items were modified to eliminate direct attribution of feeling states to alcohol consumption. At
each measurement occasion, responses to each subscale were summed to create individual
sedation (α = .89) and stimulation scores (α = .93).

Subjective intoxication. Similar to previous research (e.g., Earleywine & Erblich, 1996;
Newlin, 1985), at each assessment participants indicated how drunk they felt (“How drunk do
you feel right now?”) using a 10-point scale (1 = not drunk at all; 10 = the most drunk I’ve ever
been).

Procedure
Participants attended an initial (nondrinking) lab session where they completed the self-report measures as well as a battery of cognitive tasks germane to the aims of the larger study from which the current data were derived. One-to-three weeks later ($M = 19.1$ days) participants returned for a second (alcohol challenge) session. Participants were asked to eat a light meal 2-4 hours before their appointment. After providing informed consent, women were given a urine stream pregnancy test to self-administer (none tested positive); men were also asked to void the bladder. Participants completed a baseline BAES and subjective intoxication assessment, and then were randomly assigned to receive an active placebo (diluted [10-proof] vodka and tonic water; 0.04 g/kg ethanol), or alcohol beverage (100-proof vodka and tonic water; 0.80 g/kg ethanol for men [0.72 g/kg for women]; average peak breath alcohol concentration [BrAC] = 0.082, $SD = 0.012$). In both conditions beverages were mixed in front of participants, their contents poured from Smirnoff® vodka and Schweppes® tonic bottles, and divided into three equal-sized drinks, consumed at the rate of one every 8 min. Participants in both conditions were told that their drinks contained “a moderate amount of alcohol;” as shown in Table 2, placebo participants estimated they had consumed > 2.5 standard drinks, indicating the manipulation was effective. Total beverage was isovolumic across conditions.

After beverage consumption and following a 5-min absorption period, BrAC, BAES, and subjective intoxication measures were administered every 5-6 min until BrAC reached 0.065% for alcohol participants (or after one BrAC measurement for placebo participants), at which time the cognitive task battery was initiated. These measures were re-administered after every other cognitive task (approximately every 20 min). Upon completion of the cognitive tasks, BrAC and subjective effects were assessed every 5 min until BrAC descended from peak to 0.075%, at which time the cognitive battery was completed again; as during the ascending limb, BrAC,
BAES and subjective intoxication were assessed after every other task. Upon completion of the second round of cognitive tasks, placebo participants were debriefed and dismissed. Participants in the alcohol condition were retained in the lab until they were sober (BrAC ≤ 0.02%; see NIAAA, 2004).

Analytic Approach

The primary aims of this report involve comparisons of non-nested models (i.e., whether the ASQ or the SRE affords better prediction of a given effect). Traditional null-hypothesis significance testing (NHST) via F-ratio cannot accommodate comparison of non-nested models; therefore, model comparisons were carried out using Akaike Information Criterion (AIC; Akaike, 1974; Sakamoto et al., 1986). The AIC is an unbiased estimator of the amount of information lost in approximating a data set with a model (Burnham & Anderson, 2002). Thus, AIC provides a measure of goodness-of-fit that can be compared across several models fit to the same data (Schermelleh-Engel et al., 2003). Although formulae for AIC vary in the literature (see O’Boyle & Williams, 2011), AIC can be represented simply as:

$$\text{AIC}_i = -2\log(L_i) + 2V_i$$

where $L_i$ is the likelihood of the data given model $M_i$ and $V_i$ is the number of free parameters in model $M_i$. Lower values of AIC indicate a better fit; hence, the model with the lowest AIC is the best fitting model. The quality of any other model $M_i$ can be quantified by the difference in AIC between that model and the best-fitting model (i.e., $\Delta\text{AIC}_i$).

To perform a pairwise comparison between two models, one can convert $\Delta\text{AIC}_i$ into an evidence ratio, which gives the odds that one model provides a better fit to the data relative to the other. Ratios of less than 5:1 indicate slight evidence, ratios between 5:1 and 30:1 indicate moderate to strong evidence, and ratios in excess of 30:1 indicate very strong evidence.
(Burnham & Anderson, 2002). Model comparison through evidence ratios is straightforward and, unlike NHST, can support continuous rather than dichotomous quantification of evidence (Wagenmakers & Farrell, 2004).

Participants varied greatly in their alcohol pharmacokinetics (see Li, 2000), and due to the necessity of reaching specific BrACs to begin task sets, the number and timing of observations varied across participants. Thus, models were fit with Hierarchical Linear Modeling (HLM), which is capable of nesting repeated observations within participants and is robust to different numbers of observations per individual. HLM also can model changes in slopes over time. All models included a random intercept of subject.

Because the shape of the relationship between time and alcohol effects was expected to differ as a function of BrAC limb (e.g., Holdstock & de Wit, 1998), a regression spline was included in each model. For each individual, the spline variable at time $t$ is equal to $p_{\text{max}}(0, \text{time } t - \text{time}_{\text{peakBrAC}})$. During the ascending limb this variable is equal to 0; during the descending limb it is equal to the time elapsed since peak BrAC. This places a “knot” at the time of the individual’s peak BrAC, allowing the trajectory of the relationship between time and outcome variables to change at that time. For placebo participants, the spline was yoked to the average time when alcohol participants achieved maximum BrAC ($t = 80$ min).

To address whether ASQ scores predict inter-individual variability in post-consumption alcohol responses, as well as how the ASQ’s prediction of variation in response trajectories compares with the SRE, each subjective response outcome was modeled as a function of the interaction of time post-consumption, beverage group, sex, and one of the alcohol sensitivity measures (i.e., ASQ-Heavy, ASQ-Light, or SRE 3-mo.), as well as all lower-order interactions and main effects. This model allows for an effect of time (effects varying as BrAC rises and
falls), moderated by beverage group (alcohol participants should experience more change over
time than placebo participants), moderated by sex (men and women might differ in their
response to alcohol), moderated by scores on the ASQ or SRE (those with higher scores should
show different effect trajectories than those with lower scores). By comparing analogous ASQ-
based and SRE-based models, one can determine whether scores on one measure more
effectively capture the variance in the data than scores on the other measure.

AIC is valid for model comparison only when all models are fit to the same data.

Therefore, participants who were missing either the entire ASQ or SRE or did not provide
alcohol use data were excluded from all analyses ($n = 5$ placebo; $n = 10$ alcohol). Additionally,
individuals in the alcohol group whose peak BrAC did not reach at least 0.059% ($n = 9$) or was
greater than 0.12% ($n = 1$) were excluded, as were placebo participants who did not finish their
beverage ($n = 9$). One participant was excluded for reporting nonzero subjective intoxication at
baseline. Therefore, the final sample used for analyses included 423 individuals ($ns = 219$ and
204 in the alcohol and placebo groups, respectively).

**Results**

Bivariate correlations among primary study variables and sample demographic
characteristics are given in Table 3. AIC-based comparative fit statistics and $R^2$ estimates for
primary models are presented in Table 4. Model-derived estimated trajectories for each outcome
measure are presented in Figures 1-3.

**Stimulation**

As shown in Table 4, when predicting post-consumption stimulation ratings the model
based on ASQ-Light was strongly preferred to the SRE 3-mo. ($\Delta AIC = 10.61$; evidence ratio $=$
202) and ASQ-Heavy models ($\Delta AIC = 25.18$, evidence ratio $= 2.9 \times 10^5$). As depicted in Figure
1 and consistent with our hypotheses (see King et al., 2011), higher scores on the ASQ-Light factor predicted greater stimulation during ascending BrAC. In theory, individual differences in alcohol sensitivity should modulate subjective response only after alcohol has been consumed. Alternatively, ASQ scores could reflect alcohol-related expectancies or capture a generalized sensitivity to affective states. To test these alternatives, the best fitting (ASQ-Light) model was tested without the interaction term involving beverage group. The loss of prediction caused by dropping this interaction was dramatic ($\Delta$AIC = 19.7; evidence ratio = $1.90 \times 10^4$), indicating that the effect of ASQ-Light on stimulation ratings depends on alcohol consumption.

**Sedation**

Next, models predicting post-consumption sedation ratings were compared. Here, ASQ-Heavy produced the best-fitting model, which was strongly preferred over the SRE 3-mo. ($\Delta$AIC = 15.8, evidence ratio = 2,640) and ASQ-Light models ($\Delta$AIC = 8.23, evidence ratio = 61.3). As shown in Figure 2, relative to lower scores, higher scores on ASQ-Heavy predicted less sedation across time, and this pattern was more apparent following alcohol than following placebo consumption. Dropping the interaction with beverage group led to a substantial loss of prediction ($\Delta$AIC = 12.9; evidence ratio = 631).

**Subjective Intoxication**

Unlike both stimulation and sedation, subjective intoxication was best predicted by the SRE 3-mo. model, which was strongly preferred over the ASQ-Heavy ($\Delta$AIC = 13.6; evidence ratio = 893) and ASQ-Light models ($\Delta$AIC = 19.9; evidence ratio = 21,400). Figure 3 shows that higher scores on SRE 3-mo. were associated with lower subjective intoxication throughout the post-drinking period, and this difference was more pronounced following alcohol relative to placebo consumption. The loss of prediction caused by dropping the interaction with beverage
group was dramatic (ΔAIC = 36.9; evidence ratio = 2.05 × 10^6). A similar pattern is evident for the ASQ, but the score terciles appear to differentiate less clearly than for the SRE.

**Sensitivity versus Typical Alcohol Use**

A common concern with measures like the SRE and ASQ is that scores may simply reflect respondents’ recent alcohol involvement. If so, then models including the AlcQF variable should perform just as well as models including ASQ or SRE scores. This possibility was tested by comparing additional sets of models: (1) using the AlcQF variable in place of ASQ or SRE, and (2) using both the AlcQF and ASQ or SRE in the same model.

For stimulation effects, the ASQ-Light model and AlcQF model afforded similar prediction, with only very slight evidence in favor of ASQ-Light (ΔAIC = 1.07, evidence ratio = 1.71). However, the AlcQF model was rather strongly preferred over the SRE 3-mo. model (ΔAIC = 9.53, evidence ratio = 118). Compared to the model with AlcQF alone, the model including both AlcQF and ASQ-Light performed decidedly better (ΔAIC = 22.2, evidence ratio = 67200), as did the SRE 3-mo. model (ΔAIC = 13.7, evidence ratio = 951). The model including both AlcQF and ASQ-Light performed much better compared to an analogous model including AlcQF and SRE 3-mo. (ΔAIC = 8.52, evidence ratio = 71).

For sedation effects, the ASQ-Heavy model performed far better than the AlcQF model (ΔAIC = 29.0, evidence ratio = 1.99 × 10^6). The SRE 3-mo. model also outperformed the AlcQF model (ΔAIC = 13.2, evidence ratio = 753). Compared to the model with AlcQF alone, the model including both AlcQF and ASQ-Heavy was strongly preferred (ΔAIC = 25.5, evidence ratio > 3.4 × 10^5), as was the model including both AlcQF and SRE 3-mo. (ΔAIC = 19.3, evidence ratio = 15,900). The model including both AlcQF and ASQ-Heavy performed better than an analogous AlcQF and SRE 3-mo. model (ΔAIC = 6.18, evidence ratio = 22).
Finally, for subjective intoxication, the SRE 3-mo. model dramatically outperformed the AlcQF model ($\Delta\text{AIC} = 53.1$, evidence ratio $= 3.4 \times 10^{11}$), as did the ASQ-Heavy model ($\Delta\text{AIC} = 39.5$, evidence ratio $= 3.86 \times 10^8$). Compared to the model with AlcQF alone, the model including both AlcQF and SRE 3-mo. was very strongly preferred ($\Delta\text{AIC} = 59.0$, evidence ratio $= 6.62 \times 10^{12}$), as was the model including both AlcQF and ASQ-Heavy ($\Delta\text{AIC} = 32.8$, evidence ratio $= 1.30 \times 10^7$). The model including both AlcQF and SRE 3-mo. performed decidedly better compared to an analogous model including AlcQF and ASQ-Heavy ($\Delta\text{AIC} = 26.3$, evidence ratio $= 5.09 \times 10^5$).

**Discussion**

The SRE is the most widely-used self-report measure of alcohol sensitivity. However, the SRE’s relative utility in predicting subjective effects has never been directly tested against an alternative self-report measure. The goals of the current study were to evaluate the validity of the ASQ as such an alternative measure. We tested a family of hierarchical models in which SRE and ASQ factor scores were used to predict changes in self-reported stimulation, sedation and intoxication over time following alcohol or placebo consumption. We expected ASQ factor scores to reliably differentiate subjective responses over time for participants who consumed alcohol (but not placebo), such that higher ASQ scores (LS) would predict decreased sedation and subjective intoxication and increased stimulation. Moreover, due to its broader sampling of both stimulation- and sedation-related effects, we also expected the ASQ to provide better fit to the data relative to the SRE in predicting these outcomes.

Findings were largely consistent with these hypotheses, providing the first direct evidence for the construct validity of the ASQ. Several lines of evidence support this conclusion. First, trajectories of subjective response over time as a function of ASQ scores (see
Figures 1-3) showed that higher ASQ-Light scores were associated with greater stimulation, consistent with modified differentiator model predictions (King et al., 2011), whereas higher ASQ-Heavy scores predicted lower sedation and subjective intoxication ratings. Second, in models directly comparing the predictive utility of the ASQ and SRE, the ASQ afforded the best prediction of both stimulation and sedation. The SRE, in contrast, was better at predicting subjective intoxication (but see the additional models described in the online supporting information). Third, for each of these dependent measures responses were strongly affected by interactions involving beverage group and sensitivity scores, indicating that ASQ and SRE scores reflect sensitivity to the pharmacological effects of alcohol, as opposed to expectancy-based effects or affective fluctuations more broadly.

Incremental validity evidence for the ASQ was obtained in models comparing ASQ scores and typical alcohol use (AlcQF) as predictors of subjective effects. In each of these models, ASQ scores outperformed AlcQF and contributed substantially to model prediction over AlcQF alone, providing direct evidence that the ASQ assesses meaningful variability in alcohol sensitivity beyond what is accounted for by alcohol involvement. Interestingly, AlcQF provided better prediction of stimulation compared to the SRE, likely reflecting the lack of stimulation-related items on that measure.

The current study had numerous strengths, including a large sample, sophisticated methodological design, and the measurement of multiple domains of subjective response under both ascending and descending BrAC, but it also suffered from some limitations. Notably, participants were alone and engaged in a number of cognitive tasks over the course of the alcohol challenge session. Participants’ social isolation (see Doty & de Wit, 1995), coupled with fatigue
due to completing the cognitive tasks, could help to explain the overall low levels of stimulation
(and decline in stimulation throughout the session) observed here. An additional limitation is that alcohol response was measured only with self-report; objective measures used in some previous research (e.g., body sway, cortisol, heart rate) were not included here. However, given that the SRE was validated using self-report responses to the Subjective High Assessment Scale (SHAS; Judd et al., 1977), it is unlikely that this limitation poses a serious threat to the validity of the findings. Future research will benefit from the use of various objective measures and by querying stimulant and sedative effects that are both positive and negative (Morean et al., 2013). Also, it should be stressed that effects of initial sensitivity and effects associated with changes in sensitivity that can occur with drinking experience (i.e., tolerance) cannot be disentangled with the current data. Finally, characteristics of the current sample differed somewhat from samples used to initially validate the SRE, in that we did not conduct diagnostic interviews to exclude individuals meeting criteria for AUD, and we did exclude very light drinkers and nondrinkers.

In summary, the current study provides the first evidence that the ASQ is a reliable predictor of a variety of subjective effects of alcohol measured in the lab. Moreover, the current data suggest differing strengths for the ASQ and SRE. While both ASQ and SRE scores reflect sensitivity to pharmacological effects of alcohol beyond what is accounted for by typical alcohol use, model comparisons indicated that the ASQ outperforms the SRE in predicting post-consumption changes in stimulation and sedation but the SRE is preferred for predicting a simpler subjective intoxication index. The current results go beyond previous findings by assessing the validity of both the SRE and the ASQ in the same sample, using a statistical technique well suited for this purpose and appropriate for the structure of the data. These
findings have implications for research into the risk profile characterized by differential sensitivity to acute effects of alcohol. Wide use of this instrument will allow researchers to understand how sensitivity to both stimulant and sedative effects work dynamically to indicate risk for AUD and related problems.
References


Schuckit, M. A., Smith, T. L., Trim, R. S., Allen, R. C., Fukukura, T., Knight, E. E., et al. (2011). A prospective evaluation of how a low level of response to alcohol predicts later...


Figure Captions

Figure 1. BAES stimulation ratings across time as a function of score terciles on ASQ-Heavy (left panel), ASQ-Light (middle panel), and SRE 3-mo. (right panel) factors for participants in the placebo group (top row) and alcohol group (bottom row). ASQ-Heavy = Alcohol Sensitivity Questionnaire heavy-drinking factor; ASQ-Light = ASQ light-drinking factor; SRE 3-mo. = Self-Rating of the Effects of Alcohol Form “most recent 3-month period in which you drank” factor. BAES = Biphasic Alcohol Effects Scale. Score terciles, where “Low” represents the lower third of ASQ or SRE factor scores (i.e., HS) and “High” represents the upper third of factor scores on those measures (i.e., LS), were created for graphical representation purposes only; all analyses were carried out using continuous scores.

Figure 2. BAES sedation ratings across time as a function of score terciles on ASQ-Heavy (left panel), ASQ-Light (middle panel), and SRE 3-mo. (right panel) factors for participants in the placebo group (top row) and alcohol group (bottom row). ASQ-Heavy = Alcohol Sensitivity Questionnaire heavy-drinking factor; ASQ-Light = ASQ light-drinking factor; SRE 3-mo. = Self-Rating of the Effects of Alcohol Form “most recent 3-month period in which you drank” factor. BAES = Biphasic Alcohol Effects Scale. Score terciles, where “Low” represents the lower third of ASQ or SRE factor scores (i.e., HS) and “High” represents the upper third of factor scores on those measures (i.e., LS), were created for graphical representation purposes only; all analyses were carried out using continuous scores.

Figure 3. Subjective intoxication ratings across time as a function of score terciles on ASQ-Heavy (left panel), ASQ-Light (middle panel), and SRE 3-mo. (right panel) factors for participants in the placebo group (top row) and alcohol group (bottom row). ASQ-Heavy = Alcohol Sensitivity Questionnaire heavy-drinking factor; ASQ-Light = ASQ light-drinking factor; SRE 3-mo. = Self-Rating of the Effects of Alcohol Form “most recent 3-month period in which you drank” factor. BAES = Biphasic Alcohol Effects Scale.
factor; SRE 3-mo. = Self-Rating of the Effects of Alcohol Form “most recent 3-month period in which you drank” factor. BAES = Biphasic Alcohol Effects Scale. Score terciles, where “Low” represents the lower third of ASQ or SRE factor scores (i.e., HS) and “High” represents the upper third of factor scores on those measures (i.e., LS), were created for graphical representation purposes only; all analyses were carried out using continuous scores.
Footnotes

1. The first version of the ASQ (O’Neill et al., 2002) contained 16 items. However, in subsequent (unpublished) analyses, it was determined that one of those items, “Have you ever felt any effects from drinking alcohol?”, provided little discriminative information for identifying levels of alcohol sensitivity, and therefore it was dropped.

2. As with the SRE, scoring the ASQ begins with averaging the number of drinks a participant reports for each endorsed effect; ergo, a given item can be included in the score only if the participant reports having experienced that effect from drinking alcohol. This leads to a nonrandom pattern of missing data in which the number of endorsed items correlates with the number of drinks reported, which in turn can systematically bias sensitivity scores. See Lee et al. (2015) for scoring approaches to reduce this problem.

3. A variety of models were tested for the current report, including some in which the SRE was represented by the average of responses to all three timeframes. Model comparison results using that version of the SRE scoring were similar to those reported here. Other models used BrAC instead of time as a predictor of post-consumption subjective effects. The pattern of conclusions drawn from models using BrAC was highly similar to models using time, unsurprising given the close association between these two variables. Finally, we tested models in which the two ASQ factors were combined to form a single predictor of subjective effects. Those models are reported in Table S1 in the online supporting material.
Table 1

*Alcohol Sensitivity Questionnaire (ASQ) Items and their Factor Loadings*

<table>
<thead>
<tr>
<th>ASQ Items</th>
<th>Factor 1</th>
<th>Factor 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do you ever experience a hangover after drinking alcohol? $^H$</td>
<td>.671</td>
<td></td>
</tr>
<tr>
<td>Do you ever pass out after drinking alcohol? $^H$</td>
<td>.885</td>
<td></td>
</tr>
<tr>
<td>Do you ever throw up (vomit) after drinking alcohol? $^H$</td>
<td>.934</td>
<td></td>
</tr>
<tr>
<td>Do you ever feel nauseated after drinking alcohol? $^H$</td>
<td>.944</td>
<td></td>
</tr>
<tr>
<td>Do you ever forget part of an evening (i.e., blackouts) after drinking</td>
<td>.938</td>
<td></td>
</tr>
<tr>
<td>alcohol? $^H$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do you ever feel dizzy or feel things spinning after drinking alcohol?</td>
<td>.831</td>
<td></td>
</tr>
<tr>
<td>$^H$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do you ever become more talkative after drinking alcohol? $^L$</td>
<td>.857</td>
<td></td>
</tr>
<tr>
<td>Do you ever become more flirtatious after drinking alcohol? $^L$</td>
<td>.842</td>
<td></td>
</tr>
<tr>
<td>Do you ever feel high or “buzzed” after drinking alcohol? $^L$</td>
<td>.765</td>
<td></td>
</tr>
<tr>
<td>Do you ever feel more socially at ease after drinking alcohol? $^L$</td>
<td>.865</td>
<td></td>
</tr>
<tr>
<td>Do you ever feel more relaxed after drinking alcohol? $^L$</td>
<td>.714</td>
<td></td>
</tr>
<tr>
<td>Do you ever feel sluggish after drinking alcohol? $^L$</td>
<td>.684</td>
<td></td>
</tr>
<tr>
<td>Do you ever feel less inhibited after drinking alcohol? $^L$</td>
<td>.772</td>
<td></td>
</tr>
<tr>
<td>Do you ever feel that your driving would be affected after drinking</td>
<td>.513</td>
<td></td>
</tr>
<tr>
<td>alcohol? $^L$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do you ever feel sedated or sleepy after drinking alcohol? $^L$</td>
<td>.499</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* For each item to which respondents indicate “yes,” they are asked to respond to a follow-up question to indicate the number of drinks associated with experiencing the effect in question. Items marked with superscript “$^L$” comprise the lighter-drinking factor; follow-up questions are structured: “IF YES, what is the _minimum_ number of drinks you could consume before…” Items marked with superscript “$^H$” comprise the heavier-drinking factor; follow-up questions are structured: “IF YES, what is the _maximum_ number of drinks you could consume without…”
Table 2

*Means (and SDs) of Demographic Characteristics, Alcohol Sensitivity, Alcohol Use and Alcohol Problems Variables, and Drink Estimates as a Function of Experimental Group*

<table>
<thead>
<tr>
<th>Variables</th>
<th>Group</th>
<th></th>
<th></th>
<th>Mean comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Alcohol</td>
<td>Placebo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>23.4</td>
<td>23.2</td>
<td></td>
<td>$t(421) = -0.62, p = 0.532$</td>
</tr>
<tr>
<td>Sex (% male)</td>
<td>52%</td>
<td>46%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AlcQF</td>
<td>7.72 (6.9)</td>
<td>7.09 (7.0)</td>
<td></td>
<td>$t(421) = -0.90, p = 0.369$</td>
</tr>
<tr>
<td>Neg. Con.</td>
<td>4.39 (6.3)</td>
<td>3.84 (5.9)</td>
<td></td>
<td>$t(421) = -0.91, p = 0.361$</td>
</tr>
<tr>
<td>AUD</td>
<td>2.01 (3.3)</td>
<td>1.79 (2.8)</td>
<td></td>
<td>$t(420) = -0.88, p = 0.382$</td>
</tr>
<tr>
<td>SRE 3-mo.</td>
<td>6.29 (2.2)</td>
<td>6.07 (2.1)</td>
<td></td>
<td>$t(421) = -1.02, p = 0.301$</td>
</tr>
<tr>
<td>ASQ-Heavy</td>
<td>8.79 (2.9)</td>
<td>8.51 (2.9)</td>
<td></td>
<td>$t(421) = -1.01, p = 0.312$</td>
</tr>
<tr>
<td>ASQ-Light</td>
<td>3.43 (1.3)</td>
<td>3.35 (1.3)</td>
<td></td>
<td>$t(421) = -0.65, p = 0.516$</td>
</tr>
<tr>
<td>Estimated drinks</td>
<td>4.03 (1.3)</td>
<td>2.57 (1.4)</td>
<td></td>
<td>$t(420) = -10.56, p &lt; .001$</td>
</tr>
</tbody>
</table>

*Note.* AlcQF = quantity × frequency of alcohol use; Neg. Con. = alcohol-related negative consequences; AUD = alcohol-related negative consequences that resemble symptoms of alcohol use disorder; SRE 3-mo. = average of Self-Rating of the Effects of Alcohol form, “most recent consecutive 3-month period in which you drank” items; ASQ-Heavy = average of Alcohol Sensitivity Questionnaire heavy-drinking factor items; ASQ-Light = average of ASQ light-drinking factor items; Estimated drinks = the number of standard alcoholic drink equivalents participants believed were contained in the drinks they consumed in the lab session. For both the SRE and ASQ, means shown here represent the average number of drinks associated with the experience of queried alcohol effects. See the text for explanations of how variables reported in the table were calculated.
Table 3

*Correlations among Primary Study Variables and Sample Demographic Characteristics*

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Age</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Sex</td>
<td>.05</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. ASQ-Light</td>
<td>-.18**</td>
<td>.35**</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. ASQ-Heavy</td>
<td>-.21**</td>
<td>.46**</td>
<td>.57**</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. SRE 3-mo.</td>
<td>-.21**</td>
<td>.38**</td>
<td>.60**</td>
<td>.70**</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Max. BrAC</td>
<td>.04</td>
<td>.08</td>
<td>.02</td>
<td>.05</td>
<td>.04</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. AUD</td>
<td>-.14**</td>
<td>.02</td>
<td>.16**</td>
<td>.20**</td>
<td>.20**</td>
<td>.06</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Neg. Con.</td>
<td>-.14**</td>
<td>.03</td>
<td>.13**</td>
<td>.22**</td>
<td>.20**</td>
<td>.05</td>
<td>.94**</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>9. AlcQF</td>
<td>-.15**</td>
<td>.24**</td>
<td>.29**</td>
<td>.35**</td>
<td>.36**</td>
<td>.04</td>
<td>.48**</td>
<td>.50**</td>
<td>--</td>
</tr>
<tr>
<td>10. Binge/wk</td>
<td>-.29**</td>
<td>.03</td>
<td>.25**</td>
<td>.29**</td>
<td>.37**</td>
<td>.09*</td>
<td>.39**</td>
<td>.44**</td>
<td>.65**</td>
</tr>
</tbody>
</table>

*Note.* ASQ-Light = Alcohol Sensitivity Questionnaire, light-drinking factor score; ASQ-Heavy = ASQ heavy-drinking factor score; SRE 3-mo. = factor score from the Self-Rating of the Effects of Alcohol form, “most recent consecutive 3-month period in which you drank” items; Max. BrAC = maximum BrAC reached during the alcohol administration session (alcohol group participants only); AUD = AUD-specific alcohol-related negative consequences; Neg. Con = other alcohol-related negative consequences; AlcQF = quantity × frequency of alcohol use; Binge/wk = number of binge drinking episodes per week. See the text for explanations of how these variables were calculated.

**p < .01; *p < .05.
### Table 4

*Fit and Model Comparison Statistics for Alcohol Sensitivity Measures Predicting Post-Consumption Subjective Response Measures*

<table>
<thead>
<tr>
<th>Models</th>
<th>AIC</th>
<th>Δi</th>
<th>Marginal $R^2$</th>
<th>Conditional $R^2$</th>
<th>Evidence ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stimulation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASQ-Light</td>
<td>10443</td>
<td>0</td>
<td>.084</td>
<td>.772</td>
<td>--</td>
</tr>
<tr>
<td>ASQ-Heavy</td>
<td>10468</td>
<td>25.18</td>
<td>.088</td>
<td>.771</td>
<td>&gt;1000 : 1</td>
</tr>
<tr>
<td>SRE 3-mo.</td>
<td>10454</td>
<td>10.61</td>
<td>.086</td>
<td>.772</td>
<td>202 : 1</td>
</tr>
<tr>
<td><strong>Sedation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASQ-Light</td>
<td>11290</td>
<td>8.23</td>
<td>.097</td>
<td>.745</td>
<td>61 : 1</td>
</tr>
<tr>
<td>ASQ-Heavy</td>
<td><strong>11282</strong></td>
<td>0</td>
<td><strong>.096</strong></td>
<td><strong>.747</strong></td>
<td>--</td>
</tr>
<tr>
<td>SRE 3-mo.</td>
<td>11298</td>
<td>15.76</td>
<td>.092</td>
<td>.746</td>
<td>&gt;1000 : 1</td>
</tr>
<tr>
<td><strong>Subjective Intoxication</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASQ-Light</td>
<td>12031</td>
<td>19.94</td>
<td>.355</td>
<td>.695</td>
<td>&gt;1000 : 1</td>
</tr>
<tr>
<td>ASQ-Heavy</td>
<td>12025</td>
<td>13.59</td>
<td>.355</td>
<td>.694</td>
<td>893 : 1</td>
</tr>
<tr>
<td>SRE 3-mo.</td>
<td><strong>12011</strong></td>
<td>0</td>
<td><strong>.360</strong></td>
<td><strong>.693</strong></td>
<td>--</td>
</tr>
</tbody>
</table>

*Note.* ASQ-Light = Alcohol Sensitivity Questionnaire light-drinking factor score; ASQ-Heavy = ASQ heavy-drinking factor score; SRE 3-mo. = factor score from the Self-Rating of the Effects of Alcohol, “most recent consecutive 3-month period in which you drank” items. All models represent the Time x Beverage group x Sensitivity measure (ASQ-Light, ASQ-Heavy, or SRE 3-mo.) x Sex interaction term. AIC = Akaike Information Criteria; Δi = difference in AIC between a given model (i) and the best-fitting model within a family of models; the best-fitting model for each measure is shown in boldface. Marginal $R^2$ indicates the prediction of variance achieved through fixed effects alone; Conditional $R^2$ indicates the prediction of variance achieved through fixed and random effects (see Nakagawa & Schielzeth, 2013). Evidence ratio = odds that the model in question provides a poorer fit relative to the best-fitting model.
Figure 1

The figure displays the BAES stimulation ratings over time elapsed post-drinking (min) for different groups:

- **ASQ-Heavy**
- **ASQ-Light**
- **SRE 3-mo.**

Scores are categorized into tertiles:
- **Score tertile**
  - Low
  - Medium
  - High

The x-axis represents time elapsed post-drinking in minutes, ranging from 0 to 300. The y-axis represents BAES stimulation ratings, ranging from 0 to 40. The lines indicate the trend of ratings over time for each category.
Figure 2

![Graph showing BAES sedation ratings over time for ASQ-Heavy, ASQ-Light, and SRE 3-mo. for Placebo and Alcohol conditions, with score tertile levels indicated.]
Figure 3

Graph showing subjective intoxication ratings over time elapsed post-drinking (min) for different groups: ASQ-Heavy, ASQ-Light, and SRE 3-mo. with placebo and alcohol conditions. The graph indicates that subjective ratings peak and then decline over time, with different terciles (Low, Medium, High) represented by different line styles.
Online Supporting Information

Additional Models

The primary HLMs tested for this report compare the relative predictive utility of ASQ-Heavy, ASQ-Light and SRE 3-mo. (along with sex, time since drinking, beverage group, and their interactions) in accounting for post-consumption change in stimulation, sedation and subjective intoxication. Numerous additional models were also tested, including those in which the ASQ-Heavy and ASQ-Light factors were summed to create an additional predictor (ASQ-Sum) and those in which SRE and ASQ item means (i.e., numbers of drinks associated with the queried effects) were used instead of factor scores. Those modeling results are summarized here. Note that the base model in each case includes sex, time since drinking, and beverage group (and their interactions) along with the sensitivity measure in question.

Combining ASQ Factors

The factor analyses reported in the main text confirm a two-factor structure for the ASQ. However, it is also possible that the sum of these two factors will afford additional prediction, beyond that associated with each factor independently. This possibility was tested with additional, exploratory models in which the two ASQ factors were added to form a single ASQ-Sum predictor variable for each participant. Results of those comparisons are given in Table S1 and are summarized here.

Stimulation. As shown in Table S1, the model based on ASQ-Light remained the best predictor of post-consumption stimulation ratings, followed by the SRE 3-mo. model (ΔAIC = 10.61; evidence ratio = 202). The model based on the sum of the two ASQ factors performed considerably worse than ASQ-Light alone (ΔAIC = 14.7; evidence ratio = 1549). Thus, adding the ASQ-Sum variable did not improve prediction of changes in post-consumption stimulation over that afforded by the ASQ-Light factor alone or the SRE 3-mo. factor.

Sedation. In contrast to stimulation, post-consumption sedation was best predicted by the model based on the sum of the two ASQ factors, which was moderately preferred over the model based on the ASQ-Heavy factor alone (ΔAIC = 7.07; evidence ratio = 34) and very strongly preferred
over the model based on SRE 3-mo. ($\Delta AIC = 22.8; \text{ evidence ratio } = 9.1 \times 10^4$). Thus, adding the ASQ-Sum variable provided better prediction of changes in post-consumption sedation than either ASQ factor alone or the SRE 3-mo. factor.

**Subjective intoxication.** The primary models tested for this report, comparing both of the ASQ factors and the SRE, indicated a strong preference for the SRE in predicting subjective intoxication (see the main text for details). However, the auxiliary model comparisons showed that the model based on the ASQ-Sum variable was essentially equivalent to the model based on SRE 3-mo. ($\Delta AIC = 0.25, \text{ evidence ratio } = 1.13 \text{ in favor of ASQ-Sum}$). In addition, when sex is not included as a predictor, the model based on ASQ-Sum is strongly preferred over all other models ($\Delta AIC \geq 12.92, \text{ evidence ratio } \geq 639$). Thus, it seems that summing the two ASQ factors considerably improves the measure’s prediction of subjective intoxication, such that it becomes equivalent to the SRE in this regard.

**Using Item Means Instead of Factor Scores**

Given the rather mixed evidence of fit for the SRE 3-mo. factor reported in the main text (CFI = .95, RMSEA = .21), one could argue that the apparent advantage of the ASQ over the SRE in predicting stimulation and sedation is an artifact of the factor score approach. To examine this possibility, we conducted a set of models in which relevant item means (adjusted for missing values using standardized person mean imputation; see Lee et al., 2015) were used to represent the SRE 3-mo and ASQ (Light, Heavy, and summary score) factors.

**Stimulation.** The model based on the ASQ-Light items mean was strongly preferred to both the SRE 3-mo items mean model ($\Delta AIC = 21.9, \text{ evidence ratio } = 6 \times 10^4$) and the ASQ-Heavy items mean model ($\Delta AIC = 29.4, \text{ evidence ratio } = 2.4 \times 10^6$). These results are very much in-line with the findings using factor scores, presented in the main text.

**Sedation.** The model based on the ASQ-Heavy items mean was slightly preferred to the model based on the ASQ-Light items mean ($\Delta AIC = 2.9, \text{ evidence ratio } = 4.2$) and strongly preferred to the model based on the SRE 3-mo items mean ($\Delta AIC = 13.1, \text{ evidence ratio } = 701$). These results also replicate the findings reported using factor scores (main text).
Subjective intoxication. The model based on the ASQ-summary score mean was preferred over all other models in predicting subjective intoxication ratings. The preference ranged from strong compared to the ASQ-Heavy items mean model ($\Delta$AIC = 7.0, evidence ratio = 33) to very strong compared to the ASQ-Light items mean model ($\Delta$AIC = 8.3, evidence ratio = 62) to dramatic compared to the SRE 3-mo items mean model ($\Delta$AIC = 14.8, evidence ratio = 1608). Given that the ASQ-summary score variable is comprised of considerably more items than the SRE 3-mo mean (i.e., 15 vs. 4), we also compared the ASQ-summary score variable to a variable representing the total SRE score (i.e., all four queried alcohol effects across all three queried time periods). This total SRE score also performed dramatically worse compared to the ASQ-summary score variable in predicting subjective intoxication ($\Delta$AIC = 31, evidence ratio = $5.3 \times 10^6$).

In summary, several auxiliary models were tested for this report, using different methods of scoring the SRE and ASQ and different combinations of their items. In all cases ASQ scores were preferred over SRE scores in predicting stimulation and sedation; in some cases ASQ scores also were preferred for predicting subjective intoxication.
Table S1
*Predicting Post-Consumption Subjective Response Measures with ASQ Factors, their Combination, and the SRE*

<table>
<thead>
<tr>
<th>Models</th>
<th>AIC</th>
<th>Δi</th>
<th>Marginal $R^2$</th>
<th>Conditional $R^2$</th>
<th>Evidence ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stimulation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASQ-Light</td>
<td>10468</td>
<td>0</td>
<td>.084</td>
<td>.772</td>
<td>--</td>
</tr>
<tr>
<td>ASQ-Heavy</td>
<td>10443</td>
<td>25.18</td>
<td>.088</td>
<td>.771</td>
<td>&gt;1,000 : 1</td>
</tr>
<tr>
<td>ASQ-Sum</td>
<td>10458</td>
<td>14.71</td>
<td>.088</td>
<td>.772</td>
<td>&gt;1,000 : 1</td>
</tr>
<tr>
<td>SRE-3 mo.</td>
<td>10454</td>
<td>10.61</td>
<td>.086</td>
<td>.772</td>
<td>201.8 : 1</td>
</tr>
<tr>
<td><strong>Sedation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASQ-Light</td>
<td>11290</td>
<td>15.30</td>
<td>.097</td>
<td>.745</td>
<td>&gt;1,000 : 1</td>
</tr>
<tr>
<td>ASQ-Heavy</td>
<td>11282</td>
<td>7.07</td>
<td>.096</td>
<td>.747</td>
<td>34.3 : 1</td>
</tr>
<tr>
<td><strong>ASQ-Sum</strong></td>
<td>11275</td>
<td>0</td>
<td>.099</td>
<td>.747</td>
<td>--</td>
</tr>
<tr>
<td>SRE-3 mo.</td>
<td>11298</td>
<td>22.83</td>
<td>.092</td>
<td>.746</td>
<td>&gt;1,000 : 1</td>
</tr>
<tr>
<td><strong>Subjective Intoxication</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASQ-Light</td>
<td>12031</td>
<td>20.19</td>
<td>.355</td>
<td>.695</td>
<td>&gt;1,000 : 1</td>
</tr>
<tr>
<td>ASQ-Heavy</td>
<td>12025</td>
<td>13.84</td>
<td>.355</td>
<td>.694</td>
<td>&gt;1,000 : 1</td>
</tr>
<tr>
<td><strong>ASQ-Sum</strong></td>
<td>12011</td>
<td>0</td>
<td>.359</td>
<td>.696</td>
<td>--</td>
</tr>
<tr>
<td>SRE-3 mo.</td>
<td>12011</td>
<td>0.25</td>
<td>.360</td>
<td>.693</td>
<td>1.13 : 1</td>
</tr>
</tbody>
</table>

*Note.* All models reported here represent the Time x Beverage group x Sensitivity measure (ASQ-Light, ASQ-Heavy, ASQ-Sum or SRE) x Sex interaction term. AIC = Akaike Information Criteria; Δi = difference in AIC between a given model (i) and the best-fitting model within a family of models; the best-fitting model for each measure is shown in boldface. ASQ-Light = Alcohol Sensitivity Questionnaire light-drinking factor score; ASQ-Heavy = ASQ heavy-drinking factor score; ASQ-Sum = sum of ASQ-Light and ASQ-Heavy factors; SRE 3-mo. = factor score from the Self-Rating of the Effects of Alcohol form, “most recent consecutive 3-month period in which you drank” items. Marginal $R^2$ indicates the prediction of variance achieved through fixed effects alone; Conditional $R^2$ indicates the prediction of variance achieved through fixed and random effects. Evidence ratio = odds that the model in question provides a poorer fit relative to the best-fitting model.