Published in final edited form as:

Alcohol Clin Exp Res. 2018 June ; 42(6): 1122–1131. doi:10.1111/acer.13634.

Hangover in Post College-Aged Drinkers: Psychometric Properties of the Hangover Symptom Scale (HSS) and the Hangover Symptom Scale – Short Version (HSS-5)

Jesus Chavarria, Ph.D.¹, Sandra Y. Rueger, Ph.D.², and Andrea King, Ph.D.^{1,*}

¹Department of Psychiatry & Behavioral Neuroscience, University of Chicago, Chicago, IL 60637

²Department of Psychology, Wheaton College, Wheaton, IL 60187

Abstract

Background—Alcohol hangovers have been found to be a common and costly consequence of alcohol misuse. However, there is only limited psychometric support for instruments to accurately measure hangovers beyond the college-age years. This study investigated the psychometric properties of the Hangover Symptom Scale (HSS) and the Hangover Symptom Scale – Short Form (HSS-5) including the internal consistency, convergent and discriminant validity, construct validity, and the measurement invariance of these scales between light and heavy drinkers, individuals with a positive and negative family history of alcohol use disorder (AUD), and men and women in a post college-aged sample. The association of the HSS and HSS-5 with alcohol use problems was also tested.

Methods—Participants were 294 nonalcoholic light and heavy social drinkers (age range 21 to 35 years; 57.8% male) enrolled in the Chicago Social Drinking Project. They completed the HSS as part of a take-home packet completed outside of the laboratory. The psychometric properties of the HSS and HSS-5 were tested.

Results—Stronger psychometric support was found for the HSS-5 relative to the full HSS. While both measures demonstrated strong internal consistency reliability, convergent and discriminant validity, and an association to alcohol use problems, only the HSS-5 showed construct validity as determined by confirmatory factor analysis. Further, only the HSS-5 showed measurement invariance between men and women, light and heavy drinkers and individuals with a positive and negative family history of AUD.

Conclusions—This was the first study to examine the psychometric properties of the HSS and HSS-5 in a post college-aged sample and the first to investigate the measurement invariance of these measures as a function of sex, drinking history, and family history of AUD. This study supports the use of the HSS-5 as a reliable and valid brief measure of frequency of hangover symptoms.

*Corresponding Author: Andrea King, Ph.D., Professor, Department of Psychiatry & Behavioral Neuroscience, University of Chicago, 5841 S. Maryland Avenue (MC-3077), Chicago, IL 60637, aking@bsd.uchicago.edu., Telephone: 773-702-6181. DR. JESUS CHAVARRIA (Orcid ID: 0000-0001-5864-2899)

Keywords

hangover; alcohol use; post-college age drinkers; Hangover Symptom Scale

Introduction

Alcohol-induced hangovers consist of symptoms such as headache, fatigue, lightheadedness, dehydration, anxiety, agitation, and low mood that appear after a bout of heavy drinking when the drinker's blood alcohol concentration returns to 0.00 (Stephens et al., 2008; Verster et al., 2010). Hangovers represent the most commonly experienced negative consequence of alcohol misuse (Verster et al., 2009) with over 75% of adults reporting experiencing at least one hangover in their lifetime (Harburg et al., 1993). In terms of societal impact, hangovers are estimated to account for an annual \$148 billion USD in economic costs, mainly related to academic and workplace absenteeism and poor work performance (Ames et al., 1997; Stockwell, 1998; Ragland et al., 2002; Pidd et al., 2006; Stephens and Verster, 2010). Historically, the measurement of hangover in clinical research was hampered by a lack of instruments with demonstrated strong psychometric properties (Poikolainen, 1983; Karlsson and Romelsjö, 1997; Kauhanen et al., 1997; Maney et al., 2002; Piasecki et al., 2005; Batty et al., 2006). However, recent work has begun to better operationalize hangovers by conducting experiments to validate measures by investigating hangover symptoms the day after heavy drinking episodes (Jackson, 2008; Piasecki et al., 2010; Robertson et al., 2012; Penning et al., 2013) and after laboratory induced hangovers (Rohsenow et al., 2007) as compared to non- and/or light-drinking days.

In accordance with the Alcohol Research Group's recommendations for sound hangover research (Verster et al., 2010), several scales have been created to measure the frequency and severity of hangover symptoms. The Acute Hangover Scale (Rohsenow et al., 2007) and the Alcohol Hangover Severity Scale (Penning et al., 2013) were both tested and validated using pre-post drinking survey and laboratory induced hangover designs. The Hangover Symptom Scale (HSS; Slutske et al., 2003), a third measure that measures the frequency and severity of experienced hangovers, was one of the first multi-item scales developed to assess hangovers. The HSS retrospectively assesses past year and lifetime frequency of typical symptoms of hangover based on participants' recollection of the feelings experienced on mornings after drinking. Items include: headache, tiredness, difficulty sleeping, nausea, extreme thirst or dehydration, vomiting, trembling and shaking, difficulty concentrating, anxiety, weakness, depression, sweating, and light and sound sensitivity. The HSS has shown initial psychometric support in terms of internal consistency reliability (Cronbach's α s range = .80 to .86), item-scale correlations (r s = .35 to .68), predictive validity (Jackson, 2008; Robertson et al., 2012; Piasecki et al., 2010; Piasecki et al., 2017) and has been associated with alcohol use problems, even after controlling for the quantity of alcohol consumed and the frequency of becoming intoxicated (Slutske et al., 2003; Piasecki et al., 2010; Robertson et al., 2012; Penning et al., 2013).

While the HSS is one of the most widely used hangover scales and is recommended when the retrospective assessment of hangovers is of interest (Stephens et al., 2014), its use would

be enhanced by further evaluation of its psychometric properties in independent and novel populations. One limitation is that construct validity of the HSS using confirmatory factor analysis (CFA) has not been demonstrated. Such support would provide further evidence that the HSS accurately measures the hangover construct rather than being a face valid measure that only appears to assess for hangover symptoms (Cronbach & Meehl, 1955; DeVellis, 2016). A second limitation is that relatively few studies have used the HSS with non-college aged samples. One study followed 20–25 year olds ($M=22.4$) and found past-year hangover scores to remain relatively stable from baseline to year 1 (Piasecki et al., 2017). To our knowledge, no other studies have used the HSS to assess past-year hangover frequency in non-college aged samples. As such, there is a need for psychometric support of this measure in this older emerging adult population. Lastly, despite studies showing differences in hangover symptoms between men and women (Slutske et al., 2003; Piasecki et al., 2010; Robertson et al., 2012; Piasecki et al., 2017) and individuals with and without a positive family history of alcohol use disorder (AUD; Slutske et al., 2003; Piasecki et al., 2010; Robertson et al., 2012), no studies on the HSS have investigated its measurement invariance (i.e., determining if the number and meaning of the construct(s) being measured is the same between these groups). Tests of measurement invariance are important as they show that individuals in comparison groups are conceptualizing the construct in the same way and that the strength of the relationship between items on a scale and the construct are the same. Although not often tested, measurement invariance is a requirement when investigating differences in means across groups (Hoyle & Smith, 1994), as variance in measurement across comparison groups could lead to misleading results.

Finally, the length of the 13-item HSS may also be an issue when brevity and timeliness of test administration is of concern, as in epidemiological research or with computerized and mobile phone assessment methods. To address this issue, a Hangover Symptom Scale – Short Form with 5 items (HSS-5) was developed and tested (Slutske et al., 2003). Items of the HSS-5, including headache, tiredness, nausea, difficulty concentrating, and weakness, were chosen from a principal component analysis showing they were the best indicators of the hangover construct. A review of the literature also shows that these five symptoms appear after drinking episodes and are not due to intoxication (i.e., sleep disturbance, vomiting) or withdrawal (trembling and shaking, light and sound sensitivity; see Prat et al., 2009 for review), suggesting that they may be more purely assessing hangover symptoms. The internal consistency of this shortened scale was strong ($\alpha = .79$), the item-scale correlations for the five items ranged from 0.53 – 0.62 and the pattern of correlations between the HSS-5 and alcohol-related problems were reportedly almost identical to those with the full version (Slutske et al., 2003). Additional studies have validated the HSS-5 for hangover frequency and severity (Jackson, 2008; Piasecki et al., 2010; Jackson et al., 2013; Piasecki et al., 2017). These findings suggest that the HSS-5 is not only a viable, but potentially desirable measure of hangover symptoms. Additional psychometric support for the HSS-5 could enable greater use of this abbreviated scale.

Thus, the goal of the current study was to follow the Alcohol Research Group's recommendation (Verster et al., 2010) by further examining the psychometric properties of the HSS and HSS-5 in young adults beyond the college-age years to extend the utility of these measures. Four main research questions guided the current study: (1) Does CFA

support the hangover construct as measured by the HSS and HSS-5 in young adults beyond college-age?; (2) Do scale scores supported by the CFA have strong internal consistency reliability and convergent and discriminant validity?; (3) Are the HSS and HSS-5 associated with alcohol use problems; (4) Does the factor structure supported by the CFA show invariance between heavy and light drinkers, men and women, and individuals with and without a family history of AUD?

Materials and Methods

Participants

The sample consisted of 294 nonalcoholic light ($n = 86$) and heavy ($n = 208$) social drinkers enrolled in the Chicago Social Drinking Project between March 2004 and October 2011. Heavy drinkers included persons who consumed at least 10 but no more than 40 standard alcoholic drinks per week and engaged in regular binge drinking (e.g., binge drinking on one to five times per week, on average, as their predominant adult pattern). Light drinkers were persons who consumed at least one but no more than five standard alcoholic drinks per week and engaged in no or rare binge drinking, which typically occurred while in college. Eligibility for participation was assessed using the alcohol and other substance dependence modules of the Structured Clinical Interview for DSM-IV – nonpatient version (SCID-IV; First et al., 1995) and the Timeline Follow Back (Sobell and Sobell, 1992). Individuals with past or current major medical or psychiatric conditions including alcohol or other substance dependence were excluded from the study.

Procedure

As part of the larger Chicago Social Drinking Project (King et al., 2011) participants completed a take-home packet of surveys, including the HSS. The study was fully approved by the University of Chicago Institutional Review Board.

Measures

Hangover Symptom Scale (HSS; Slutske et al., 2003)—The HSS is a 13-item Likert-type scale that assesses the frequency of experiencing hangover symptoms over the past year the morning after drinking alcohol (0 [Never] = 0% of the time, 1 [Occasionally] = about 25% of the time, 2 [About half of the time] = 50% of the time, 3 [Most of the time] = 75% of the time, and 4 [Every time I drink alcohol] = 100% of the time). As stated earlier, the 13 items include headache, tiredness, difficulty sleeping, nausea, extreme thirst/dehydration, vomiting, trembling/shaking, difficulty concentrating, anxiety, weakness, depression, sweating, and light/sound sensitivity. Scores range from 0–52. For the HSS-5, five items from the HSS (headache, tiredness, nausea, difficulty concentrating, and weakness) were scored in an identical fashion to the HSS, with scores ranging from 0–20. The HSS was chosen it was the only valid and reliable scale of hangover at the start of the Chicago Social Drinking Project (e.g., 2004).

Quantity Frequency Interview (QFI; Cahalan, 1969)—The QFI assesses subjects' reported consumption of wine, beer, and liquor over the previous 6 months. Standardized drinks were as follows: beer = 12oz, liquor = 1.5oz, wine = 5oz. These standardized drinks

were then converted to ounces of 100% ethanol consumed per day. This measure has been found to be a reliable and valid assessment of quantity and frequency of alcohol consumption (Cahalan, 1969).

Alcohol Use Disorder Identification Test (AUDIT; Saunders et al., 1993)—The AUDIT is a Likert-type scale that screens for hazardous or harmful drinking. Scores range from 0–40 with scores ≥ 8 denoting probable hazardous drinking. The AUDIT has been found to be a reliable measure in determining individuals with hazardous drinking patterns (Barry and Fleming, 1993; Saunders et al., 1993.).

Drinker Inventory of Consequences-Recent Consequences Form (DrInC-2R; Miller et al., 1995)—The DrInC-2R is a 45-item Likert-type scale that assesses for the frequency of experienced negative drinking consequences over the past three months with scores ranging from 0–80. This measure assesses consequences related to drinking in the following domains: physical (physical health harmed by drinking), interpersonal (family or friends complain about my drinking), intrapersonal (felt bad about myself because of my drinking), impulse control (driven a motor vehicle after 3+ drinks), and social responsibility (I have gotten in trouble because of drinking). The measure has been found to be a valid and reliable instrument for assessing alcohol use problems (Miller et al., 1995; Cisler and Zweben, 1999).

Structured Clinical Interview for DSM-IV (SCID-IV; First et al., 2002)—The SCID-IV is a structured clinical interview that assesses the diagnostic criteria for lifetime DSM-IV alcohol abuse and dependence. Subjects' symptoms rating could be coded as 1 (no evidence of symptom), 2 (symptom met at subthreshold level), or 3 (symptom completely met). A past “heaviest drinking year” AUD symptom count, as defined by DSM-5, was calculated by adding up all symptoms that were met completely (i.e., all symptoms rated as a 3) with AUD symptom counts ranging from 0–11. This method was chosen to align with the codification system of the DSM-5 that emphasizes that AUD occurs on a continuum. The DSM-IV criteria are similar to those in the DSM-5 with the exception of the item on alcohol-related legal problems that was dropped in the DSM-5, and the inclusion of a craving item that had not yet been developed (American Psychiatric Association, 2013; Hasin et al., 2013). These SCID-IV modules have been found to be reliable and valid assessment of alcohol use disorders (Kidorf et al., 1998; First et al., 2002).

Family History-Research Diagnostic Criteria (Andreasen et al., 1977)—The FH-RDC is a two-generational biological family history tree measure used to assess family history of AUD. The participant is asked to indicate which family members had alcohol use problems. Potential family members included the participant's paternal and maternal grandparents, aunts, and uncles and the participant's parents and siblings. Follow up questions concerning the specific type of problems related to alcohol use (e.g., legal, health, marital/family, occupational, other [e.g., fighting, losing friends], and receiving treatment) were asked about all family members who were identified as having alcohol problems. Family history positive (42%) participants were defined as those reporting at least one primary biological relative or at least two secondary biological relatives with AUD, family

history negative (36%) participants had no AUD in biological relatives within the past two generations. Participants who were unable to provide sufficient information to be categorized as either positive or negative were grouped as family history unknown (22%). Family history unknown data were used in all analyses except the measurement invariance testing between individuals with and without a family history of AUD. The FH-RDC has been found to be a reliable and valid measure for assessing family history of AUD (Andreasen et al., 1977).

Data Analytic Strategy

The underlying factor structure of the HSS and HSS-5 was tested using CFA with maximum likelihood estimator in Mplus Version 8.0 (Muthen and Muthen, 1998–2017). Following the recommendation of Hu and Bentler (1998), fit indices used included the comparative fit index (CFI) and standardized root mean square residual (SRMR). A CFI .90 or greater and a SRMR less than .10 were used as indicators of acceptable fit and a CFI greater than .95 and a SRMR less than .08 were used to indicate good fit (Hu and Bentler, 1998).

Next, scale scores were calculated using the items supported in the CFA, and Cronbach's alphas were calculated in SPSS 24.0 (IBM Corporation, 2016). A Cronbach's alpha above .70 is considered adequate internal consistency reliability (DeVellis, 2016), which is a necessary, but not sufficient, indicator of a unidimensional construct (Streiner, 2003; Tavakol and Dennick, 2011). Convergent validity was examined by correlating a single item from the DrINC-2R that assesses hangover directly (e.g., "I have had a hangover or felt bad after drinking") with the HSS and HSS-5 total scores. In addition, five items from the DrINC-2R that appeared to assess physical symptoms of hangover (e.g., "Because of my drinking, I have not eaten properly" and "I have had a hangover or felt bad after drinking"), and possible hangover consequences (e.g., "I have missed days of school or work because of drinking;" "The quality of my work has suffered because of my drinking;" and "I have failed to do what is expected of me because of my drinking") were combined into a scale score and regressed on the HSS and HSS-5, separately, while controlling for the quantity and frequency of alcohol use. Convergent validity, evidenced by a correlation or beta equal or greater than .45, determines if the construct in question is highly correlated to a related construct (De Von et al., 2007). Lastly, discriminant validity was investigated by regressing the interpersonal, intrapersonal, and impulse control subfactors, which did not include any items of the hangover related items used in the previous analyses, on HSS and HSS-5, separately. Discriminant validity, evidenced by a correlation or beta less than .45, determines if the construct is poorly correlated with an unrelated construct (De Von et al., 2007).

The association between the HSS hangover constructs and the alcohol use problem variables (e.g., AUDIT Total score, DrINC-2R total score [minus hangover items used in the previous analyses], SCID-AUD symptom count) was tested using six linear regression models. Next SEM was used to test the associations between latent hangover (HSS & HSS-5) and a latent alcohol use problems construct. In these models, a latent Alcohol Use Problems construct comprising the total scores from the AUDIT, DrINC-2R (minus hangover items used in previous analyses), and SCID-AUD symptom count was regressed on the HSS and HSS-5 Hangover construct separately while controlling for the frequency and quantity of alcohol

use by using the QFI manifest variable. The latent Hangover constructs comprised the 13 and 5 items of the HSS and HSS-5, respectively.

To examine the invariance of factor structure for drinking group, sex, and family history, SEM was used following the procedures discussed by Bryne (2012) and Vandenberg & Lance (2000). According to this method, a sequential pattern of more restrictive models are compared. A test of configural invariance (i.e. same basic factor structure across groups with no constraints across item loadings, intercepts, or residual variances) was first tested. The test of configural invariance determines if the hangover construct was conceptualized in the same way in both groups. Next, tests of metric invariance (i.e., item loadings constrained, but intercepts and residual variances allowed to vary) and scalar invariance (i.e., item loadings and intercepts constrained, but residual variances allowed to vary) were run. Metric invariance indicates that the strength of the relationship between each item and the hangover construct are similar in both groups being tested. Scalar invariance indicates that individuals in each group would score the same on the hangover construct regardless of group membership. Improvement in fit between the unconstrained and constrained models was tested for statistical significance by computing the differences in chi-square (χ^2) and degrees of freedom from the two models being compared. A non-significant χ^2 difference test indicates no significant improvement in the more restrictive, constrained model, suggesting invariance of measurement at the level tested.

Results

Descriptive Statistics

The sample was 58% male with an age range between 21 and 35 years (mean 25.6 ± 3.0 SD years). The majority of the sample was between 23–29 years of age ($n=212$; 72%), 9% ($n=25$) were between 30–35 years of age, and 19% ($n=57$) were between 21–22 years of age. Racial breakdown of the sample was as follows: 77% Caucasian, 11% African American, 6% Asian American, <1% American Indian/Alaskan Native, and 6% Multiracial. In terms of ethnic composition, 9% of participants identified as Hispanic/Latino.

Comparisons of light and heavy drinkers showed they did not differ on any of these demographic characteristics (p 's=ns), with the exception of age (26.08 ± 3.41 for light drinkers vs. 25.08 ± 2.69 for heavy drinkers; $p = .02$). As expected, heavy drinkers had higher HSS and HSS-5 scores relative to light drinkers (HSS: 11.6 ± 6.2 for heavy drinkers vs. 5.4 ± 5.7 for light drinkers; HSS-5: 5.8 ± 3.2 vs. 3.0 ± 3.3 , respectively; p 's<.001).

Endorsement of the HSS items ranged from 21% to 89% (see Table 1). Investigation of the correlations among the entire sample showed a significant correlation between the HSS and HSS-5 ($r=.93$, $p<.001$). Additionally, the HSS and HSS-5 were each correlated with all of the alcohol related variables (see Table 2). The correlations for all of the demographic, hangover, and alcohol related measures are provided on Table 2.

Confirmatory Factor Analysis

HSS—The results of the CFA of the HSS for the full sample did not support a one-factor model. Although the SRMR of .05 suggested adequate fit, the CFI of .88 indicated

inadequate fit of the one factor model to the data. Further, as shown in Figure 1, nearly half of the HSS items had weak standardized loadings, i.e., below .60 (Guadagnoli and Velicer, 1988). In spite of the less than adequate fit of the model, Cronbach's alpha was calculated, and was found to be strong ($\alpha = .87$).

HSS-5—The results of the CFA of the HSS-5 for the full sample supported a one-factor model. The SRMR of .02 and the CFI of .99 both indicated excellent fit of the one-factor model to the data. Further, as shown in Figure 2, the factor loadings for all items were $> .60$, which further suggests good model fit. Cronbach's alpha was calculated with these five items, and the internal consistency of the HSS-5 was found to be strong ($\alpha = .83$).

Convergent & Discriminant Validity

HSS—Supporting the convergent validity of the HSS, a significant correlation was found between the HSS and the single hangover item of the DrInC-2R ($r = .59, p < .001$). Additionally, a linear regression analysis showed the HSS was significantly associated with the combination of the DrInC-2R hangover items ($\beta = .52, p < .001$) after accounting for the quantity and frequency of drinking alcohol ($\beta = .29, p < .001$). With regards to discriminant validity, linear regressions between the HSS and the interpersonal ($\beta = .40, p < .001$), intrapersonal ($\beta = .43, p < .001$), and impulse control ($\beta = .40, p < .001$) subfactors of the DrInC-2R were all less than the .45 criterion (De Von et al., 2007).

HSS-5—A significant correlation between the HSS-5 and the single hangover item of the DrInC-2R ($r = .59, p < .001$) supported the convergent validity of the HSS-5. A linear regression also found a significant association between the HSS-5 and the combined hangover items ($\beta = .50, p < .001$) after controlling for the quantity and frequency of alcohol consumption ($\beta = .31, p < .001$). With regards to discriminant validity, associations between the HSS-5 and the interpersonal ($\beta = .38, p < .001$), intrapersonal ($\beta = .44, p < .001$), and impulse control ($\beta = .33, p < .001$) subfactors of the DrInC-2R were less than the .45 threshold (De Von et al., 2007).

Association between Hangover and Alcohol Use Problems

HSS—The linear regression analyses showed the HSS was significantly associated with AUDIT ($\beta = .48, p < .001$), DrInC-2R ($\beta = .51, p < .001$), and SCID-AUD symptom count ($\beta = .44, p < .001$). Next, SEM to test a full model including both alcohol use and hangover experiences predicting alcohol use problems showed mixed results with a SRMR of .07 indicating good fit, but a CFI of .89 indicating poor fit. Despite the mixed model fit results, significant effects were found for both the Hangover construct ($\beta = .52, p < .001$) and the QFI ($\beta = .59, p < .001$) on the Alcohol Use Problems construct.

HSS-5—The linear regressions showed the HSS-5 was also significantly associated with AUDIT ($\beta = .43, p < .001$), DrInC-2R ($\beta = .47, p < .001$), and SCID-AUD symptom count ($\beta = .39, p < .001$). The SEM testing a full model including both alcohol use and hangover experiences predicting alcohol use problems showed good model fit (SRMR = .07, CFI = .97) and significant effects for the Hangover construct and QFI on Alcohol Use Problems (see Figure 3).

Tests of Invariance

HSS—Results did not support measurement invariance for the HSS for drinking group, sex, or family history of AUD. Specifically, the SRMR showed adequate fit for the configural model for drinking history, sex, and family history of AUD, but the CFI indicated poor fit (see Table 3). As configural invariance is a necessary condition for further tests of invariance, tests of metric and scalar invariance were not conducted (Vandenberg & Lance, 2000).

HSS-5—In contrast, for the HSS-5, results supported full invariance of measurement for drinking group and family history of AUD as evidenced by good fit of the configural, metric, and scalar models and non-significant χ^2 difference tests between the most and least restrictive models (see Table 3). With regards to sex, there was support for partial measurement invariance with the configural and metric models showing good fit to the data (Table 3); however, the comparison of the metric model to the configural model resulted in a significant χ^2 difference test ($\chi^2 [df] = 10.80 [4], p = .02$), suggesting that at least one item loading was not equivalent between men and women. A closer examination of the loadings showed that the *difficulty concentrating* item displayed the greatest item loading difference between men and women, and the constraint on this item was relaxed. The fit indices for the metric model with the one item allowed to vary indicated close fit, and a comparison of this modified metric model and the configural model resulted in a non-significant χ^2 difference test (see Table 3), indicating partial metric invariance. The scalar model, with the *difficulty concentrating* item-loading allowed to vary between men and women, showed acceptable fit to the data (see Table 3). The comparison of the scalar model with the modified metric model also resulted in a non-significant significant χ^2 difference test (see Table 3). These results show that the HSS-5 is partially invariant between men and women, indicating that any differences in scores between men and women could be attributed to either differences in the Hangover construct or the difference in the way the difficulty concentrating item is related to the Hangover construct between men and women.

Discussion

As the prevalence and consequences of heavy drinking continues well past college ages (i.e., 18–22; Perkins, 1999; Merline et al., 2004; Center for Behavioral Health Statistics and Quality, 2016), we sought to conduct an independent test of the psychometric properties of the HSS, as well as the shorter HSS-5, in post college-aged drinkers. Consistent with previous research, we found strong internal consistency among the HSS items, convergent validity through the HSS's association with hangover-related items of other alcohol use problem scales, and an association between the HSS and alcohol use problems (Slutske, et al., 2003; Robertson et al., 2012; Penning et al., 2013). Discriminant validity was also found with associations less than the .45 criterion between the HSS and alcohol related interpersonal, intrapersonal, and impulse control problems, as compared with hangover-related items. However, the HSS did not demonstrate an underlying factor structure that supported a single hangover construct. In addition, the HSS did not show invariance across drinking group, sex, or family history of AUD. Thus, evidence of construct validity was weak. In contrast, the shorter HSS-5 showed strong psychometric properties including

internal consistency and strong construct validity. Specifically, the HSS-5 demonstrated convergent and discriminant validity and an underlying factor structure that supported a single hangover construct with full measurement invariance across drinking group and family history of AUD. Only partial measurement invariance across sex was found for the HSS-5.

It is unclear why the full HSS did not show support for stronger construct validity: Despite showing convergent and discriminant validity, three items of the HSS with the lowest factor loadings were sweating, difficulty sleeping, and vomiting, which are all symptoms in the DSM-IV and DSM-5's description of alcohol withdrawal. Since hangover and withdrawal represent different constructs and physiological consequences of heavy drinking, these particular items may be assessing a withdrawal subfactor as opposed to an overall hangover factor (see Prat et al., 2009 for review). Further, as withdrawal and hangover are both consequences of alcohol over consumption, they both should be strongly related to hangover-related outcomes (i.e., high correlation between hangover items of DrINC-2R), but should not fit the theoretical construct of a hangover.

This is consistent with findings in the original psychometric paper on the HSS (Slutske et al., 2003) that found two eigenvalues above 1.0, which indicates the presence of two possible factors. However, a one factor model was chosen rather than a two factor model as the second factor accounted for a small amount of variance and relatively low endorsement of the five items that constituted the second factor (Slutske et al., 2003). As the original psychometric study (Slutske et al., 2003) consisted primarily of college-aged drinkers (i.e., 18–22 years) who rarely show alcohol withdrawal (Knight et al., 2002), there may have been too infrequent of endorsement of such items to warrant a withdrawal subfactor. College drinking environments can promote atmospheres for heavy binge drinking which can induce symptoms such as vomiting (i.e., drinking games, fraternity/sorority pledging, learning personal drinking limits), but the other demands of college may not be conducive to the style of drinking necessary for physical dependence. In the current study, all HSS items, with exception of vomiting, were endorsed at a commensurate or greater rate than the Slutske et al (2003) study and numerous items also did not load strongly onto the hangover construct. Continued research of hangover in heavier drinking and older samples may enable more in-depth examination of the possibility of multiple subfactors within the HSS.

To our knowledge this study represents the first to use CFA in supporting the factor structure and measurement invariance of the HSS-5. Specifically, invariance across drinking group and family history of AUD indicated that any differences in hangover scores between light and heavy drinkers or individuals with and without a family history of AUD were solely due to differences in the hangover construct (i.e., differences in HSS-5 scores were due to differences in the frequency of experiencing hangover as opposed to group differences in the interpretation of the hangover items). Partial measurement invariance for sex suggested that differences in hangover scores is likely due to differences in the hangover construct, but may also be due to the fact that *difficulty concentrating* was a stronger indicator of hangover for women vs. men in the current sample. Overall, the study results provide support for the use of the HSS-5, a more parsimonious measure than the HSS, in post college-aged drinkers.

There were several strengths to this study, including the diversity of the sample with regards to drinking history and family history of AUD, and a moderate-to-large sample size which allowed us to test the factor structure and measurement invariance. However, several limitations are worth noting to guide further research. First, outside the HSS, no other psychometrically sound hangover surveys were employed so tests of criterion-related validity could not be conducted. Second, the sample included individuals in their 20s and 30s so while findings extend those in college-aged samples, they may not generalize to middle-aged and older populations. Third, we used symptom count for DSM-IV criteria as this study originated prior to the development of the DSM-5. While symptom count pertains to DSM-5 AUD classifications, we do not feel that symptom count of DSM-IV abuse and dependence criteria are markedly different as 10 of the 11 DSM-IV items match the DSM-5. Finally, as participants were non-alcohol dependent, it is not possible to generalize the findings to individuals with more severe drinking patterns and those with a severe AUD. Future studies will need to investigate the psychometric properties of the HSS and HSS-5 in older individuals and those with severe AUD to determine if the psychometric properties outlined in this study generalize to those subpopulations.

In sum, the current study provides psychometric support for the HSS-5 in post college-aged drinkers as the scale showed good internal consistency reliability, construct validity, convergent and discriminant validity, and measurement invariance across three different participant subgroups (sex, drinking history and family history). Consistent with the guidelines developed for hangover research (Verster et al., 2010), this study provides further evidence for the use of HSS-5 as a hangover measurement tool and extends its use for measuring hangover in post college-age drinkers and for measuring differences between men and women, light and heavy drinkers, and family history of AUD positive and negative individuals. Future studies should continue to investigate the underlying factor structure of the HSS and the possibility of a withdrawal subfactor. Additionally, researchers will need to replicate our findings in order to further strengthen the psychometric support of the HSS-5.

Acknowledgments

Appreciation is extended to Patrick McNamara and Patrick Smith for data collection and database assistance. Funding: R01-AA013746 (AK) & 3R01AA013746-14S1.

References

- Ames GM, Grube JW, Moore RS. The relationship of drinking and hangovers to workplace problems: an empirical study. *Journal of Studies on Alcohol*. 1997; 58:37–47. [PubMed: 8979212]
- Andreasen NC, Endicott J, Spitzer RL, Winokur G. The family history method using diagnostic criteria: reliability and validity. *Archives of General Psychiatry*. 1977; 34:1229–1235. [PubMed: 911222]
- Association AP. *Diagnostic and Statistical Manual*. 5. American Psychiatric Publishing; 2013.
- Barry KL, Fleming MF. The Alcohol Use Disorder Identification Test (AUDIT) and the SMAST 13: predictive validity in rural primary care setting. *Alcohol and Alcoholism*. 1993; 28:33–42. [PubMed: 8471085]
- Batty GD, Deary IJ, Macintyre S. Childhood IQ and life course socioeconomic position in relation to alcohol induced hangovers in adulthood: the Aberdeen children of the 1950s study. *Journal of Epidemiology & Community Health*. 2006; 60:872–874. [PubMed: 16973534]

- Byrne, BM. Structural Equation Modeling: Basic Concepts, Applications, and Programming. Routledge; New York, NY: 2012.
- Cahalan D. A multivariate analysis of the correlates of drinking-related problems in a community study. *Social Problems*. 1969; 17:234–247.
- Center for Behavioral Health Statistics and Quality. Key substance use and mental health indicators in the United States: results from the 2015 National Survey on Drug Use and Health. 2016. (HHS Publication No. SMA 16-4984, NSDUH Series H-51). Retrieved from <http://www.samhsa.gov/data/>
- Cisler RA, Zweben A. Development of a composite measure for assessing alcohol treatment outcome: operationalization and validation. *Alcoholism: Clinical and Experimental Research*. 1999; 23:263–271.
- Cronbach LJ, Meehl PE. Construct validity in psychological tests. *Psychological Bulletin*. 1955; 52:281–302. [PubMed: 13245896]
- DeVellis, RF. Scale development: Theory and applications. Vol. 26. Sage publications; 2016.
- DeVon HA, Block ME, Moyle-Wright P, Ernst DM, Hayden SJ, Lazzara DJ, Savoy SM, Kostas-Polston E. A psychometric toolbox for testing validity and reliability. *Journal of Nursing Scholarship*. 2007; 39:155–164. [PubMed: 17535316]
- Hasin DS, O'Brien CP, Auriacombe M, Borges G, Bucholz K, Budney A, Compton WM, Crowley T, Ling W, Petry NM, Schuckit M, Grant BF. DSM-5 criteria for substance use disorders: recommendations and rationale. *American Journal of Psychiatry*. 2013; 170:834–851. [PubMed: 23903334]
- First, MB., Spitzer, RL., Gibbon, M., Williams, JB. Structured Clinical Interview for DSM-IV-TR Axis I Disorders. Research Version, Patient Edition, SCID-I/P. New York: Biometrics Research, New York State Psychiatric Institute; 2002. Structured Clinical Interview for DSM-IV-TR Axis I Disorders. Research Version, Patient Edition, in Series November 2002.
- First MB, Spitzer RL, Gibbon M, Williams JBW, Davies M, Borus J, Howes MJ, Kane J, Pope HG, Rounsaville B. The Structured Clinical Interview for DSM-III-R Personality Disorders (SCID-II). Part II: multi-site test-retest reliability study. *Journal of Personality Disorders*. 1995; 9:92–104.
- Guadagnoli E, Velicer WF. Relation to sample size to the stability of component patterns. *Psychological Bulletin*. 1988; 103:265. [PubMed: 3363047]
- Harburg E, Gunn R, Gleiberman L, DiFranceisco W, Schork A. Psychosocial factors, alcohol use, and hangover signs among social drinkers: a reappraisal. *Journal of Clinical Epidemiology*. 1993; 46:413–422. [PubMed: 8501466]
- Hasin DS, Stinson FS, Ogburn E, Grant BF. Prevalence, correlates, disability, and comorbidity of dsm-iv alcohol abuse and dependence in the united states: results from the national epidemiologic survey on alcohol and related conditions. *Archives of General Psychiatry*. 2007; 64:830–842. [PubMed: 17606817]
- Hoyle RH, Smith GT. Formulating clinical research hypotheses as structural equation models: a conceptual overview. *Journal of Consulting and Clinical Psychology*. 1994; 62:429–440. [PubMed: 8063970]
- Hu, L-t, Bentler, PM. Fit indices in covariance structure modeling: sensitivity to underparameterized model misspecification. *Psychological Methods*. 1998; 3:424–453.
- Jackson KM. Heavy episodic drinking: Determining the predictive utility of five or more drinks. *Psychology of Addictive Behaviors*. 2008; 22:68–77. [PubMed: 18298232]
- Jackson KM, Rohsenow DJ, Piasecki TM, Howland J, Richardson AE. Role of tobacco mmoking in hangover symptoms among university students. *Journal of Studies on Alcohol and Drugs*. 2013; 74:41–49. [PubMed: 23200149]
- Karlsson G, Romelsjo A. A longitudinal study of social, psychological and behavioural factors associated with drunken driving and public drunkenness. *Addiction*. 1997; 92:447–457. [PubMed: 9177066]
- Kauhanen J, Kaplan GA, Goldberg DD, Cohen RD, Lakka TA, Salonen JT. Frequent hangovers and cardiovascular mortality in middle-aged men. *Epidemiology*. 1997:310–314. [PubMed: 9115028]
- Kidorf M, Brooner RK, King VL, Stoller KB, Wertz J. Predictive validity of cocaine, sedative, and alcohol dependence diagnoses. *Journal of Consulting and Clinical Psychology*. 1998; 66:168–173. [PubMed: 9489271]

- Kline, RB. Principles and Practice of Structural Equation Modeling: Third Edition. Guilford Press; New York: NY: 2011.
- Knight JR, Wechsler H, Kuo M, Seibring M, Weitzman ER, Schuckit MA. Alcohol abuse and dependence among US college students. *Journal of Studies on Alcohol*. 2002; 63:263–270. [PubMed: 12086126]
- Maney DW, Higham-Gardill D, Mahoney BS. The alcohol-related psychosocial and behavioral risks of a nationally representative sample of adolescents. *Journal of School Health*. 2002; 72:157–163. [PubMed: 12029813]
- Merline AC, O'Malley PM, Schulenberg JE, Bachman JG, Johnston LD. Substance use among adults 35 years of age: prevalence, adulthood predictors, and impact of adolescent substance use. *American Journal of Public Health*. 2004; 94:96–102. [PubMed: 14713705]
- Miller, WR., Tonigan, JS., Longabaugh, R. Project MATCH monograph series 4. 1995. The Drinker Inventory of Consequences (DrInC).
- Muthen, LK., Muthen, BO. Mplus User's Guide. Muthen & Muthen; Los Angeles, CA: 1998–2017.
- Penning R, McKinney A, Bus LD, Olivier B, Slot K, Verster JC. Measurement of alcohol hangover severity: development of the Alcohol Hangover Severity Scale (AHSS). *Psychopharmacology*. 2013; 225:803–810. [PubMed: 23007602]
- Perkins HW. Stress-motivated drinking in collegiate and postcollegiate young adulthood: life course and gender patterns. *Journal of Studies on Alcohol*. 1999; 60:219–227. [PubMed: 10091960]
- Piasecki TM, Sher KJ, Slutske WS, Jackson KM. Hangover frequency and risk for alcohol use disorders: evidence from a longitudinal high-risk study. *Journal of Abnormal Psychology*. 2005; 114:223. [PubMed: 15869353]
- Piasecki TM, Slutske WS, Wood PK, Hunt-Carter EE. Frequency and correlates of diary-measured hangoverlike experiences in a college sample. *Psychology of Addictive Behaviors*. 2010; 24:163. [PubMed: 20307125]
- Piasecki TM, Trela CJ, Mermelstein RJ. Hangover symptoms, heavy episodic drinking, and depression in young adults: A cross-lagged analysis. *Journal of Studies on Alcohol and Drugs*. 2017; 78:580–587. [PubMed: 28728640]
- Pidd KJ, Berry JG, Roche AM, Harrison JE. Estimating the cost of alcohol-related absenteeism in the Australian workforce: the importance of consumption patterns. *Med J Aust*. 2006; 185:637–641. [PubMed: 17181511]
- Poikolainen K. Inebriation and mortality. *International Journal of Epidemiology*. 1983; 12:151–155. [PubMed: 6874208]
- Prat G, Adan A, Sánchez-Turet M. Alcohol hangover: a critical review of explanatory factors. *Human Psychopharmacology: Clinical and Experimental*. 2009; 24:259–267. [PubMed: 19347842]
- Ragland DR, Krause N, Greiner BA, Holman BL, Fisher JM, Cunradi CB. Alcohol consumption and incidence of workers' compensation claims: a 5-year prospective study of urban transit operators. *Alcoholism: Clinical and Experimental Research*. 2002; 26:1388–1394.
- Robertson BM, Piasecki TM, Slutske WS, Wood PK, Sher KJ, Shiffman S, Heath AC. Validity of the Hangover Symptoms Scale: evidence from an electronic diary study. *Alcoholism: clinical and experimental research*. 2012; 36(1):171–177.
- Rohsenow DJ, Howland J, Minsky SJ, Greece J, Almeida A, Roehrs TA. The Acute Hangover Scale: A new measure of immediate hangover symptoms. *Addictive Behaviors*. 2007; 32:1314–1320. [PubMed: 17097819]
- Saunders JB, Aasland OG, Babor TF, De La Fuente JR, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption-II. *Addiction*. 1993; 88:791–804. [PubMed: 8329970]
- Slutske WS, Piasecki TM, Hunt-Carter EE. Development and initial validation of the Hangover Symptoms Scale: prevalence and correlates of hangover symptoms in college students. *Alcoholism: Clinical and Experimental Research*. 2003; 27:1442–1450.
- Sobell, LC., Sobell, MB. Timeline Follow-Back. In: Litten, RZ., Allen, JP., editors. *Measuring Alcohol Consumption: Psychosocial and Biochemical Methods*. Humana Press; Totowa, NJ: 1992. p. 41–72.

- SPSS Statistics for Windows [computer program]. Version Version 24.0. Armonk, NY: IBM Corp; 2016.
- Stephens R, Grange JA, Jones K, Owen L. A critical analysis of alcohol hangover research methodology for surveys or studies of effects on cognition. *Psychopharmacology*. 2014; 231:2223–2236. [PubMed: 24633471]
- Stephens R, Ling J, Heffernan TM, Heather N, Jones K. A review of the literature on the cognitive effects of alcohol hangover. *Alcohol & Alcoholism*. 2008; 43:163–170. [PubMed: 18238851]
- Stephens R, Verster CJ. Editorial [The Importance of Raising the Profile of Alcohol Hangover Research]. *Current Drug Abuse Reviews*. 2010; 3:64–67. [PubMed: 20822482]
- Stockwell T. Towards guidelines for low-risk drinking: quantifying the short- and long-term costs of hazardous alcohol consumption. *Alcoholism: Clinical and Experimental Research*. 1998; 22:63S–69S.
- Streiner DL. Starting at the beginning: an introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment*. 2003; 80:99–103. [PubMed: 12584072]
- Tavakol M, Dennick R. Making sense of Cronbach's alpha. *International Journal of Medical Education*. 2011; 2:53–55. [PubMed: 28029643]
- Vandenberg RJ, Lance CE. A Review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*. 2000; 3:4–70.
- Verster JC, van Herwijnen J, Olivier B, Kahler CW. Validation of the Dutch version of the brief young adult alcohol consequences questionnaire (B-YAACQ). *Addictive behaviors*. 2009; 34:411–414. [PubMed: 17942242]
- Verster JC, Stephens R, Penning R, Rohsenow D, McGeary J, Levy D, McKinney A, Finnigan F, Piasecki TM, Adan A, Batty GD, Fliervoet LAL, Heffernan T, Howland J, Kim D-J, Darren Kruisselbrink L, Ling J, McGregor N, Murphy RJL, van Nuland M, Oudelaar M, Parkes A, Prat G, Reed N, Slutske WS, Smith G, Young M. The alcohol hangover research group consensus statement on best practice in alcohol hangover research. *Current Drug Abuse Reviews*. 2010; 3:116–126. [PubMed: 20712593]

Figure 1. Hangover Symptom Scale with Standardized Item Loadings. HSS = Hangover Symptom Scale. All paths are significant $p < .001$.

Figure 2. Hangover Symptom Scale –Short Scale with Standardized Item Loadings. HSS-5 = Hangover Symptom Scale – Short Scale. All paths are significant $p < .001$.

Figure 3.

Structural Equation Model of Hangover on Alcohol Use Problems. QFI = Quantity Frequency Index. HSS-5 = Hangover Symptom Scale – Short Scale. AUDIT = Alcohol Use Disorder Identification Test. DrInC-2R=Drinker Inventory of Consequences (minus hangover-related items). SCID=Structured Clinical Interview of DSM-IV Disorders alcohol abuse and dependence modules total symptom count. Path loadings are standardized values. All paths are significant $p < .001$.

Table 1

Hangover Symptom Scale Item Endorsement

Item	Past Year Endorsement	Mean (Standard Deviation) ^a
Headache *	79%	1.19 (.89)
Tiredness *	84%	1.46 (1.04)
Difficulty Sleeping	31%	.40 (.72)
Nausea *	52%	.64 (.72)
Extreme Thirst/Dehydration	89%	1.91 (1.21)
Vomiting	38%	.40 (.54)
Trembling/Shaking	21%	.24 (.52)
Difficulty Concentrating *	61%	.89 (.93)
Anxiety	28%	.35 (.66)
Weakness *	62%	.82 (.85)
Depression	24%	.28 (.56)
Sweating	27%	.40 (.80)
Light/Sound Sensitivity	51%	.78 (.97)

Note:

^aMean and Standard Deviation of HSS Likert-type scale responses (0 = 0% of the time; 1 = about 25% of the time; 2 = about 50% of the time; 4 = about 75% of the time; 5 = 100% of the time).

* HSS-5 Items.

Table 2

Correlation Table of Study Variables

	1	2	3	4	5	6	7	8
1. Sex	-	-	-	-	-	-	-	-
2. Group	-.09	-	-	-	-	-	-	-
3. FH+	.04	.06	-	-	-	-	-	-
4. HSS	.09	.42***	.14**	-	-	-	-	-
5. HSS-5	.10*	.38***	.13**	.93***	-	-	-	-
6. QFI	-.25***	.61***	.10*	.21***	.18***	-	-	-
7. AUDIT	-.14**	.76***	.15***	.48***	.43***	.61***	-	-
8. DrInC-2R	-.07	.53***	.15***	.54***	.50***	.47***	.65***	-
9. SCID	-.08	.54***	.11*	.44***	.39***	.43***	.65***	.65***

Note.

*** <.01,

** <.05,

* <.10.

Sex coded as 1=male 2=female. Group=drinking group. Group coded as 1=light drinkers 2=heavy drinkers. FH+=family history of alcohol use disorder. FH+ coded as 0=FH negative, 0 = FH unknown, 1=FH positive. HSS=Hangover Symptom Scale. HSS-5=Hangover Symptom Scale-Short Scale. QFI = Quantity Frequency Index. AUDIT=Alcohol Use Disorder Identification Test. DrInC-2R=Drinker Inventory of Consequences. SCID=Structured Clinical Interview of DSM-IV Disorders alcohol abuse and dependence module symptom count.

Table 3

Measurement Invariance Tests between Drinking Groups, Family History of having an Alcohol Use Disorder, and Sex

	χ^2 (df)	p	CFI	SRMR
HSS				
Drinking Level				
Configural	-	-	.83	.07
Metric	-	-	-	-
Scalar	-	-	-	--
Family History of Alcohol Use Disorder				
Configural	-	-	.86	.07
Metric	-	-	-	-
Scalar	-	-	-	-
Sex				
Configural	-	-	.87	.06
Metric	-	-	-	-
Scalar	-	-	-	-
HSS-5				
Drinking Level				
Configural	-	-	.99	.03
Metric	6.75 (4)	.15	.98	.05
Scalar	7.09 (4)	.13	.98	.05
Family History of Alcohol Use Disorder				
Configural	-	-	.96	.04
Metric	5.27 (4)	.26	.96	.06
Scalar	6.77 (4)	.15	.95	.06
Sex				
Configural	-	-	.96	.04
Metric ^a	6.87 (3)	.08	.95	.06
Scalar ^a	8.38 (4)	.08	.94	.06

Note. χ^2 =chi-square. df=degrees of freedom. χ^2 =change in χ^2 . CFI=Comparative Fit Index. SRMR=Standard Root Mean Square Residual. HSS=Hangover Symptom Scale. HSS-5=Hangover Symptom Scale-Short Scale.

^adifficulty concentrating item-loading allowed to vary between groups.