HHS Public Access
Author manuscript
Addiction. Author manuscript; available in PMC 2018 August 01.

Published in final edited form as:
Addiction. 2017 August ; 112(8): 1495–1507. doi:10.1111/add.13830.

Alcohol Use from Adolescence through Early Adulthood: An Assessment of Measurement Invariance by Age and Gender

Jessica N. Fish[1], **Amanda M. Pollitt**[2], **John E. Schulenberg**[3], and **Stephen T. Russell**[4]

[1]Population Research Center, University of Texas at Austin, 305 E. 23rd St., Stop G1800, Austin, TX 78712

[2]Norton School of Family and Consumer Sciences, The University of Arizona, 650 N. Park Ave., Tucson, AZ 85721-0078

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

[4]Human Development and Family Sciences, University of Texas at Austin, 108 E. Dean Keeton St., Stop A2702, Austin, TX, 78712

Abstract

Background and aims—Studies on alcohol use and related constructs rarely test for measurement invariance to assess the reliability and validity of measures of alcohol use across different subpopulations of interest or ages. This failure to consider measurement invariance may result in biased parameter estimates and inferences. This study aimed to test measurement invariance of alcohol use across gender and age using a US-nationally representative sample to inform future longitudinal studies assessing alcohol use.

Design—The National Longitudinal Study of Adolescent to Adult Health, a school-based, nationally-representative longitudinal study conducted in 1994–1995, 2001–2002, and 2008.

Setting—All regions within the United States; participants were selected via a clustered sample design from 80 high-schools that represented the national population.

Participants—Youth and young adults ages 13 to 31 who had valid data on all three alcohol items within wave: 18,923 from Wave 1; 14,315 from Wave 3; and 14,785 from Wave 4.

Measurements—Alcohol use measurement models were constructed using past-year general drinking frequency, heavy episodic drinking frequency, and average quantity when drinking. Configural (factor structure), metric (factor loadings), and scalar (item intercepts) measurement invariance models were tested by gender and for each year of age from 13 to 31.

Findings—All models passed the threshold for configural invariance. Comparisons between males and females demonstrated metric (and usually scalar) non-invariance for most ages beyond middle adolescence. Nearly all one- and two-year contrasts passed metric invariance. Scalar non-

Corresponding Author, Jessica N. Fish, jessica.fish@utexas.edu, Ph: 512-471-4270.
The authors have no conflict of interest to report at this time.

invariance was most prevalent in age comparisons between late adolescence and early adulthood, particularly for tests using two-year age increments.

Conclusions—Studies that do not account for the effects of gender and age on the measurement of alcohol use may be statistically biased.

Keywords

Alcohol use; Measurement invariance; Development; Methods; Latent variable modeling

Introduction

Developmental perspectives are necessary to understand the etiology and progression of alcohol use across the lifespan [1–3]. Longitudinal studies yield important findings regarding trends and critical periods for excessive alcohol use and behavioral precursors of alcohol use disorders (AUDs) [4]. Notwithstanding their contribution, many longitudinal studies use the same alcohol use measure across subgroups of interest (e.g., males and females) or developmental periods (e.g., adolescence, early adulthood) without testing for measurement invariance (i.e., whether measures reflect the same scale and metric across groups or developmental periods; c.f. [5,6]). Yet, alcohol use initiation, escalation, and desistence vary across specific subpopulations (e.g., by gender, race/ethnicity, sexual orientation, college attendance) [7–9] from adolescence to adulthood [1,10,11]. Parameter estimates may be biased if measures are assumed to be invariant when, in fact, they are not [12–14]. Developmentally framed studies on alcohol consumption provide critical evidence for policy, prevention, and intervention efforts regarding risky alcohol use behaviors and related consequences. It is imperative that this evidence come from studies that verify that measures of alcohol use are consistent across ages and key groups of interest.

Alcohol Use across the Transition to Adulthood and Considerations for Measurement

Initiation and escalation of drinking behaviors typically occur in mid-adolescence with a peak in risky drinking during early adulthood and a subsequent decline in alcohol consumption for those transitioning into social roles and formal institutions that discourage excessive alcohol use (e.g., family and workforce) [11,15]. Particularly in the US, alcohol consumption is linked to adolescent development experiences of independence and experimentation [1] and risky alcohol use is an often expected and anticipated behavior for young adults [10,16]. For example, alcohol consumption follows a steep gradient from early to late adolescence: Past 30-day drinking doubles between 8^{th} (i.e., 13–14 years olds) and 10^{th} (i.e., 15–16 years olds) grade and 4.1% of eighth graders report HED in the previous two-weeks compared to 12.6% of 10^{th} graders and 19.4% of 12^{th} graders (i.e., 18 year olds) [17]. Thus, the construct of “alcohol use” changes in its relative importance and meaning at different points in the lifespan [15].

In addition to age-graded trends in overall alcohol use, research shows distinct change in the occurrence and extent of specific alcohol use behaviors [18,19]—frequency of alcohol use, heavy episodic drinking (HED), and average quantity when drinking—which point to potential maturational influences. Studies also show differential rates of change for specific alcohol use behaviors; thus, drinking behaviors shift over time. For instance, adolescents who drink report more consistency in the number of days they drink and the number of days they engage in HED compared to individuals of legal drinking age [20]. Thus, the presence, magnitude, and covariance of alcohol use behaviors may also vary as youth age.

Along with age differences, specific subpopulations—defined by gender, race/ethnicity, and college attendance, among others—vary in trajectories of alcohol use across the transition to adulthood [7,8,21]. Studies have shown unique developmental trends between males and females. Early adolescent females demonstrate higher rates of alcohol use and HED than their male peers [17,22]. As they age, however, young males report accelerating levels of alcohol use and HED that eventually surpass female reports on or around the age of 15 [7,17,23]. Although both males and females exhibit an increase in alcohol consumption during the transition to adulthood, males report a greater increase in frequency and amount of alcohol use, resulting in a gender disparity in alcohol use and AUDs by early adulthood [23–25]. During the gradual decline of alcohol use for those entering their late 20s and early 30s, women report a greater decrease in consumption and risky drinking compared to men and are more likely to abstain from alcohol altogether [7,25].

Epidemiological and psychosocial studies often operationalize alcohol use by combining self-report items that reflect composites of specific behaviors [26–28], such as drinking frequency, HED, average quantity when drinking, drunkenness, alcohol use disorder symptomology, and consequences associated with drinking, among others. Yet, gender and age differences in alcohol use behaviors suggest that composite scores may differentially characterize drinking for men and women or for people at different ages. This type of measurement error can alter the conclusions drawn from investigations of alcohol use. One way to investigate whether measures systematically differ across groups or age is through measurement invariance testing.

Latent Variable Modeling and Measurement Invariance

Latent variable modeling (also known as structural equation modeling) can be used to account for and attenuate the effect of measurement error that may bias parameter estimates [12,29]. This approach accommodates complex statistical models that include many items, measures, or outcomes, leading to a reduction in Type I error. However, many researchers do not conduct tests of measurement invariance which can introduce systematic error or bias [14,30].

Broadly, tests of measurement invariance assess whether observable items are consistently represented by the same underlying latent construct for different subgroups within a population [29] (see supplement for more information). Analyses that do not confirm measurement invariance may have biased estimates, jeopardizing the accuracy of results [12,14,30]. Conversely, conclusions from findings are robust when conditions of measurement invariance are met [14]. Importantly, the confirmation of invariance extends beyond latent modeling frameworks because it also substantiates the utility, or lack thereof, of non-latent (i.e., manifest) scales compiled from observed items [12,31].

There are typically four steps of measurement invariance assessment. Each step imposes an increasing number of restrictions to test whether model parameter equality constraints reveal differences across groups via a decrease in the quality of model fit [12]. First, an initial unconstrained or freely estimated configural model is examined to assess whether the factor structure (the number and pattern of factor loadings) is equivalent across groups (see Figure 1). Second, metric invariance is assessed by constraining the factor loadings to be equal across comparisons. With metric invariance, the expected magnitude of change on each item is the same across groups for every one-unit change in the latent variable. Third, scalar invariance imposes equality constraints on the intercepts of observed indicators, which examines whether mean level differences in the latent construct equivalently characterize change in the observed indicators. The fourth step, residual factorial invariance, denotes equivalence in the degree of item-specific and random error. Consistent with recent commentary, we do not test for residual factorial invariance because it is unlikely that random error is equivalent across groups or time. [12].

Measurement Invariance and Alcohol Epidemiology

There remain limited studies on the equivalence of measures across groups and developmental periods in alcohol epidemiology, including the DSM-IV criteria for alcohol dependence across clinical and non-clinical samples [32], the Alcohol Use Disorders Identification Test across gender (AUDIT) [33], problem drinking across Hispanic and white youth [34], and changes in reasons for limiting or abstaining from drinking across time [35]. In testing the psychometric properties of the Drinking Motives Questionnaire (DMQ), Cooper [36] noted factor loading differences by gender for coping motives, and differences by age for conformity motives and cautioned researchers to consider potential bias when analyzing these constructs. Not surprisingly, this study spurred numerous subsequent investigations of measurement invariance in studies using the DMQ [37–42].

Yet, even fewer studies explicitly examine the measurement of alcohol consumption independent of related antecedents and consequences of alcohol use. Bullers and colleagues [43] tested the equivalence of an alcohol use construct composed of drinking frequency, HED, and average quantity when drinking and noted that HED and average drinking quantity were more strongly related to the latent construct across groups defined by gender (male vs. female), age (≤ 40 vs. 41+), and race (Black vs. other). In a more recent study [44], findings supported the measurement invariance of alcohol use—operationalized via alcohol use frequency, HED, average number of drinks, and intoxication—across gender in an age-homogeneous college sample. Generally, the lack of studies assessing the measurement invariance of alcohol use prevents researchers from drawing definitive conclusions about the utility of such measures. However, despite implications for a better understanding of age and group differences, researchers do not typically evaluate longitudinal or group measurement invariance when testing hypotheses in observed or latent modeling frameworks (c.f. [45]). As such, we have limited understanding of how multi-item assessments of alcohol use (manifest or latent) operate in studies that assess the relative stability or change in alcohol use as people age. Such an investigation is warranted given the methodological and substantive knowledge gained from assessments of measurement invariance.

The Current Study

Testing the measurement invariance of alcohol use measures is needed given the implications of measurement bias in studies assessing the antecedents and consequences of alcohol use, especially in light of age and gender differences in alcohol use. To address this, we use a US representative, population-based sample to test the measurement structure and reliability of an NIAAA-recommended measure of alcohol use [26]. We test configural, metric, and scalar invariance during a developmentally sensitive period for alcohol use initiation, acceleration, and desistance [46] for (1) males and females within each year of age and (2) one- and two-year age group differences for males and females separately. Findings will elucidate the consistency and validity of alcohol use measures from adolescence through early adulthood across and within gender, informing methodological approaches for developmental studies of alcohol use.

Methods

Data Source and Sample

We use data from the National Longitudinal Study of Adolescent to Adult Health [47], one of the most comprehensive studies of US adolescents as they transition to adulthood. Wave 1 data were collected from adolescents enrolled in middle and high schools during the 1994–1995 school year (n = 20,745). Wave 3 occurred approximate 6 years later, when participants were between the ages of 18–24 (n = 15,197) and Wave 4 was conducted in 2008, when participants were between the ages of 24 and 32 (n = 15,701). Our sample included all participants in Wave 1, 3, and 4 who were assigned a sampling weight at baseline (n = 18,924) and provided a response to at least one of the three alcohol items at each measurement occasion. The data were restructured so that measurement models were compared by age-year and not wave, allowing us to explore specific developmental differences in alcohol use [7].

Measures

Alcohol use—As suggested by NIAAA [26], alcohol consumption was measured with three items assessing 12-month retrospective reports of drinking frequency, HED frequency, and average quantity when drinking. General drinking frequency was assessed by asking participants “During the past 12 months, on how many days did you drink alcohol?” with response options of: *never* = 0; *1 or 2 days in the past 12 months* = 1; *3–12 times in the past 12 months* = 2; *once a month or less* = 3; *2 or 3 days a month* = 4; *1 or 2 days a week* = 5; *3 to 5 days a week* = 6; and *every day or almost every day* = 7. Respondents reported HED frequency on the same scale in response to the question, “Over the past 12 months, on how many days did you drink five or more drinks in a row?” for Waves 1 and 3. In Wave 4, the item assessed the frequency of HED at “5 or more” for males and “4 or more” for females. Participants reported their average drink quantity by providing a number in response to the question, “Think of all of the times you have had a drink during the past 12 months. How many drinks did you usually have each time (A “drink” is a glass of wine, a can of beer, a wine cooler, a shot glass of liquor, or a mixed drink.)”. Answers were recoded to reflect 1 drink increments (*1 drink* = 1; *2 drinks* = 2, etc.) until *10 or more drinks* = 10. People who reported not drinking in the previous year were coded as 0 for this item.

Gender—Respondents reported their gender at Wave 1 as either *male* = 0 or *female* = 1.

Analytic Approach

Data management was performed in Stata 14.2 [48] and measurement invariance analyses were conducted in Mplus 7.4 [49]. All analyses presented are weighted and adjusted to account for Add Health's survey design [50] and missing data were accounted for via full information maximum likelihood (FIML) during latent variable estimation. Measurement invariance was established by first assessing fit of models that freely estimated parameters (configural invariance) followed by a sequence of more systematically constrained models to assess the variability of factor loadings (metric invariance) and intercepts (scalar invariance) through degradation in overall model fit [12] (see Figure 1). We used the change in CFI (i.e., ΔCFI) as our threshold for invariance [51,52]—where values greater than or equal to .010 indicate a significant decrease in overall model fit (the large sample sizes made chi-square difference tests less effective). During estimation, we constrained the variance of the latent factor to 1.0 and freely estimated factor loadings to assess item-level differences [12].

We first assessed measurement invariance between males and females for each year of age. We then assessed differences by age for males and females, independently. Age comparisons were denoted by one and two-year age groups. If models failed metric or scalar invariance, we independently assessed the loadings or intercepts of each item to identify which indicator(s) contributed to measurement non-invariance for a given comparison.

Results

Table 1 displays the number of males and females assessed and the average observed score for alcohol use for each year of age. Tests of configural invariance held for both males and females and across all age groups. That is, within both males and females, the three observed items of alcohol use formed a single factor, with each item loading highly on the single factor for all years of age.

Measurement Invariance between Males and Females by Age

Omnibus tests of measurement invariance by gender are presented in the left column of Table 2. Male and female measurement models were metric invariant (equivalent loadings but not intercepts) for ages 17–20, but only configurally invariant (nonequivalent loadings and intercepts) across gender during ages 13, 23–27, and again at 30–31. Scalar invariance (equivalent loadings and intercepts) was confirmed for males and females from ages 14–16, 19, 21–22 and again at ages 28 and 29.

Indicator factor loadings, intercepts, and residual variances for each male- and female-by-year comparison are presented in Table 3. A ΔCFI ≤ .010 for the item indicates a significant improvement in model fit when the item is freely estimated, signaling that the item differs across groups. Item level gender comparisons (see far right column of Table 3) revealed higher factor loadings for males, compared to females, for drinking frequency at the age of 13. There were also significantly higher drinking frequency and HED intercepts for males at age 17 and 20 and higher drinking frequency intercepts at age 18. At ages 23–24, 26–27 and 30–31, factor loadings for drinking frequency were higher for males; at age 25, these factor loadings were lower for males. At ages 23, 24, and 31, loadings for HED were higher for males; and at ages 25 and 26, loadings for average quantity were higher for males.

Measurement Invariance by One-Year Age Increments

Results presented on the right side of Table 2 show that with one exception all one-year age assessments confirmed metric invariance within males and females. Among males, all comparisons met criteria for scalar invariance except between 18 and 19-year-olds, which met criteria for metric invariance. Item level measurement invariance assessments (Table 3) indicated lower intercepts for drinking frequency and higher intercepts for average alcohol quantity for 18 compared to 19-year-old males.

Among females, there was scalar non-invariance for 24 vs. 25-year-olds and metric non-invariance for 30 vs. 31-year-olds (the one exception of metric non-invariance across age) (Table 2). Item level differences (Table 3) indicated that HED drinking intercepts were lower, but average quantity intercepts were higher, for 24 compared to 25-year-olds. Factor loadings were larger for HED frequency and smaller for average quantity for 30 compared to 31-year-olds.

Measurement Invariance by Two-Year Age Increments

Assessments of measurement invariance across two-year age groups for males and females are presented in Table 4. All two-year age comparisons passed metric invariance. Scalar invariance assessments failed for 19–20 vs. 21–22-year-olds among both males and females, as well as for 17–18 vs. 19–20-year-old males and 23–24 vs. 25–26-year-old females. Follow up assessments at the item level (Table 5) showed that drinking frequency and average alcohol quantity intercepts were lower for 17–18 compared to 19–20-year olds. Drinking frequency intercepts were lower, and alcohol quantity intercepts higher, for 19–20 compared to 21–22 males. For females, drinking frequency intercepts were lower at ages 19–20 than 21–22-year-olds. Intercepts for HED were lower and average alcohol quantity higher for 23–24-year-old compared to 25–26-year-olds.

Discussion

We tested configural, metric, and scalar measurement invariance of an NIAAA-recommended alcohol use measure across gender and age during the transition to adulthood. Findings suggest reasonably consistent invariance in the construct validity and utility of a composite alcohol use measure by age from 13 to 31, but only sometimes by gender. These findings extend the literature in two important ways. First, the lack of metric and scalar invariance across gender, indicating the lack of equivalent associations between observed items and item intercepts of the latent construct, provides evidence that the multi-item alcohol use measure is not necessarily comparable for males and females for most ages. That is, observed items are not equivalently associated with, nor provide a common scale for, the alcohol use latent construct for males and females. Second, findings regarding age-graded trends in (non)invariance both support and problematize previous studies that have assessed alcohol use behaviors across broad age groups. Together, findings suggest that measurement of the global construct of “alcohol use” varies as a function of gender and, to a lesser extent, age.

Measurement (Non)Invariance and Gender

The lack of construct comparability across gender is consistent with most alcohol use research demonstrating that, compared to women, men drink in qualitatively and quantitatively different ways [25]; these differences are largely attributable to gendered psychosocial and cultural factors [53]. Results show that our measure of alcohol use operates differently both in its metric and scale for males and females at different ages. The differential co-occurrence of drinking behaviors—alcohol use frequency, HED frequency, and average quantity when drinking—are distinctly associated with the latent construct of alcohol use for males and females. That is, this multi-item alcohol use scale measures items differently across genders. Consequently, this measurement bias can lead to inaccurate conclusions when estimating antecedents and consequences of alcohol use. Take stress as a predictor of alcohol use, for example; if a one-unit change in the latent construct reflects a larger change in alcohol frequency for women, but a larger change in HED frequency for men, then a positive association between stress and alcohol use reflects an increase in different behaviors for women and men (i.e., that stress is more strongly related to drinking frequency for women and heavy drinking for men) [12].

Our results show relatively little scalar and, more importantly, metric invariance between men and women after the age of 21; thus, the inclusion of both men and women in models assessing alcohol use appears to be especially problematic for studies of adults. For studies of youth, specifically, our results suggest that researchers may find that models of alcohol use demonstrate comparability across males and females given the relative consistency in loadings and intercepts during early to middle adolescence. However, this equivalency in scalar invariance dissipates during late adolescence and the transition to adulthood, though metric invariance remains through age 21. Thus, measurement invariance by gender should be checked explicitly rather than assumed, especially with older adolescents and adults.

Measurement (Non)Invariance and Age

Scalar non-invariance in models among males and females across age is not surprising considering literature on changes in the degree to which people engage in specific alcohol use behaviors from adolescence to adulthood [10,15,17]. Again, the presence of metric invariance suggests that the latent construct accurately captures changes in the (co)variance of alcohol use behaviors as people transition to adulthood. Findings regarding non-invariant intercepts, however, demonstrate that the estimated scale of the latent variable changes across age and may over- or underestimate factor scores depending on the directionality and extent of the mean level difference. This bias could misspecify the strength and association of risk factors at different ages, important information for timely and targeted prevention and intervention strategies. Points of scalar non-invariance also illuminate developmental periods when drinking behaviors change [7,17]. Differences, for example, were most prevalent in the two-year comparisons during the transition from late adolescence to the early 20s—adjacent the age of legal consumption and a time readily known for youth experimentation and acceleration in alcohol use [1].

Recommendations

We echo previous recommendations and emphasize the importance of invariance testing procedures in measurement validation and model estimation [12,31,54]. Findings here lead us to conclude that models testing correlates and outcomes related to alcohol use across gender and/or developmental periods can inadvertently introduce measurement error when unexamined or unaddressed. If tests of invariance fail at the metric or scalar level, partial measurement invariance techniques is a potentialsolution. Testing partial invariance requires the assessment of model fit across sequentially constrained models: fit is assessed while imposing equality constraints on invariant parameters while those that are non-invariant are allowed to vary freely. There is, however, no broad consensus regarding the magnitude of acceptable partial invariance. Steenkamp and Baumgartner [55] argue the need for at least one item (other than the scaling item) to be invariant at the metric and scalar level to provide stability for meaningful comparisons; however, these applications require theoretical and substantive grounding [54]. Our findings suggest that freeing a single item loading and/or intercept—typically the parameter(s) that contribute the most overall change in model fit [12]—would meet this minimum requirement for most ages, and to a lesser extent, for gender. Therefore, researchers should explore partial measurement invariance procedures in developmental analyses to mitigate measurement error and improve the interpretation of findings in light of bias.

Although many researchers attempt to address the influence of gender by including it as a covariate, this may not adequately account for the measurement error introduced when both the (co)variance and mean of items are non-invariant. Researchers could instead stratify analyses by gender, depending on research design. However, for research questions that explicitly test gender differences, measurement invariance testing is a necessary minimum to understand how alcohol use assessments may vary for males and females and consequently over- or underestimate differences or associations with other variables of interest. Partial measurement invariance may reduce error during estimation when there is minimal metric or scalar non-invariance, but our results suggest that, especially for adults, multi-item alcohol use measures vary in function and form by gender.

Finally, even strict invariant models do not provide error-proof estimation; all models, at best, provide only a utilitarian approximation of reality [56]. Therefore, knowledge of measurement invariance is a barometer from which researchers can evaluate the usefulness and limitations of measures to inform the implementation of analytic strategies that attempt to minimize error during estimation. We encourage researchers to further investigate these differences, the examinations of which could reveal important between- and within-group differences in the developmental trends of alcohol use behaviors—and substance use more generally—across adolescence and adulthood.

These contributions should be considered in the context of limitations. First, we provide an age-graded assessment of measurement invariance during the transition to adulthood given that this is a particularly vulnerable time for the development of excessive alcohol use and alcohol use disorders [46]. However, we were unable to examine the measurement invariance of alcohol use in middle and late adulthood, another important age period for the development of alcohol use disorders [1,57,58]. Future studies should test measurement invariance by age beyond the early thirties to illuminate sources of error and differences in alcohol use behavior across the lifespan. Second, although our three-item alcohol use measure reflects recommendations from NIAAA [26], there are many other commonly used measures of alcohol use that would benefit from such an assessment. Third, the time between Waves 1, 3, and 4 of the Add Health data averaged 6 years. Ideally, we would have assessed the within-person rates of change across 1 and 2 year age periods, but this was not possible with the Add Health data structure. Data that measure alcohol use across more incremental time frames, such as annually, may provide more nuanced and informed assessments of alcohol use behaviors across the transition to adulthood. Finally, this study is a methodological example of how alcohol use, as a construct, varies across developmental stages and does not necessarily reflect the many individual and contextual variables that might influence developmental and gendered differences in alcohol use, such as socioeconomic status, race, ethnicity, sexual orientation, and community level factors such as the availability of alcohol [59–62].

Future investigations of measurement invariance in developmental, etiological, and epidemiological studies could uncover important group differences that elucidate risk and strengthen understanding of alcohol use across the lifespan. Researchers should consider measurement invariance testing a standard part of their preliminary data procedures—particularly when measuring alcohol and other substance use over time.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (http://www.cpc.unc.edu/addhealth). No direct support was received from grant P01-HD31921 for this analysis.

Funding: This study was funded in part by the National Institute on Alcohol Abuse and Alcoholism (awarded to Fish) grant number F32AA023138. This research was also supported by grant, R24HD042849, awarded to the Population Research Center at The University of Texas at Austin by the Eunice Kennedy Shriver National Institute of Child Health and Human Development. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Preliminary analyses were supported by the National Institute on Alcohol Abuse and Alcoholism (awarded to Russell) grant number R01AA020270, and support for Russell from the Priscilla Pond Flawn Endowment at the University of Texas at Austin. Schulenberg acknowledges support from National Institute on Drug Abuse grants R01DA001411 and R01DA016575.

References

1. Brown SA, McGue M, Maggs J, Schulenberg J, Hingson R, Swartzwelder S, et al. A developmental perspective on alcohol and youths 16 to 20 years of age. Pediatrics. 2008; 121:S290–S310. [PubMed: 18381495]
2. Merline A, Jager J, Schulenberg JE. Adolescent risk factors for adult alcohol use and abuse: stability and change of predictive value across early and middle adulthood. Addiction. 2008; 103:84–99. [PubMed: 18426542]

3. Schulenberg, J., Patrick, ME., Maslowsky, J., Maggs, JL. The epidemiology and etiology of adolescent substance use in developmental perspective. In: Lewis, M., Rudolph, KD., editors. Handbook of developmental psychopathology. New York: Springer; 2014.
4. Schulenberg JE, Maggs JL. Destiny matters: distal developmental influences on adult alcohol use and abuse. Addiction. 2008; 103:1–6.
5. Leigh BC, Stacy AW. Alcohol expectancies and drinking in different age groups. Addiction. 2004; 99:215–27. [PubMed: 14756714]
6. Sher KJ, Wood MD, Wood PK, Raskin G. Alcohol outcome expectancies and alcohol use: a latent variable cross-lagged panel study. J Abnorm Psychol. 1996; 105:561–74. [PubMed: 8952189]
7. Chen P, Jacobson KC. Developmental trajectories of substance use from early adolescence to young adulthood: gender and racial/ethnic differences. J Adolesc Health. 2012; 50:154–63. [PubMed: 22265111]
8. Chen P, Jacobson KC. Longitudinal relationships between college education and patterns of heavy drinking: a comparison between Caucasians and African-Americans. J Adolesc Health. 2013; 53:356–62. [PubMed: 23707401]
9. Dermody SS, Marshal MP, Cheong J, Burton C, Hughes T, Aranda F, et al. Longitudinal disparities of hazardous drinking between sexual minority and heterosexual individuals from adolescence to young adulthood. J Youth Adolesc. 2014; 43:30–9. [PubMed: 23325141]
10. Schulenberg JE, Maggs JL. A developmental perspective on alcohol use and heavy drinking during adolescence and the transition to young adulthood. J Stud Alcohol Drugs. 2002; (Suppl):54–70.
11. O'Malley PM. Maturing out of problematic alcohol use. Alcohol Res Health. 2005; 28:202–5.
12. Little, TD. Longitudinal structural equation modeling. New York, NY: Guilford Press; 2013.
13. Little, TD., Card, NA., Preacher, KJ., McConnell, E. Modeling longitudinal data from research on adolescence. In: Lerner, R., Steinberg, L., editors. Handbook of adolescent psychology. Hoboken, NJ: Wiley; 2009.
14. Widaman KF, Ferrer E, Conger RD. Factorial invariance within longitudinal structural equation models: measuring the same construct across time. Child Dev Perspect. 2010; 4:10–18. [PubMed: 20369028]
15. Patrick ME, Schulenberg JE. How trajectories of reasons for alcohol use relate to trajectories of binge drinking: national panel data spanning late adolescence to early adulthood. Dev Psychol. 2011; 47:311–7. [PubMed: 21219061]
16. Neighbors C, Lee CM, Lewis MA, Fosso N, Larimer ME. Are social norms the best predictor of outcomes among heavy-drinking college students? J Stud Alcohol Drugs. 2007; 68:556–65. [PubMed: 17568961]
17. Miech, RA., Johnston, LD., O'Malley, PM., Bachman, JG., Schulenberg, JE. Volume I, Secondary school students. Ann Arbor, MI: Institute for Social Research, The University of Michigan; 2015. Monitoring the Future national survey results on drug use, 1975–2014.
18. Jackson KM, Sher KJ. Similarities and differences of longitudinal phenotypes across alternate indices of alcohol involvement: a methodological comparison of trajectory approaches. Psychol Addict Behav. 2005; 19:339–51. [PubMed: 16366806]
19. Talley AE, Sher KJ, Littlefield AK. Sexual orientation and substance use trajectories in emerging adulthood. Addiction. 2010; 105:1235–45. [PubMed: 20491728]
20. Windle M. Drinking over the lifespan: focus on early adolescents and youth. Alcohol Res Curr Rev. 2016; 38:95–101.
21. Timberlake DS, Hopfer CJ, Rhee SH, Friedman NP, Haberstick BC, Lessem JM, et al. College attendance and its effect on drinking behaviors in a longitudinal study of adolescents. Alcohol Clin Exp Res. 2007; 31:1020–30. [PubMed: 17403064]
22. Duncan SC, Duncan TE, Strycker LA. Alcohol use from ages 9 to 16: a cohort-sequential latent growth model. Drug Alcohol Depend. 2006; 81:71–81. [PubMed: 16006054]
23. Jackson KM, Sher KJ, Cooper ML, Wood PK. Adolescent alcohol and tobacco use: onset, persistence and trajectories of use across two samples. Addiction. 2002; 97:517–31. [PubMed: 12033653]

24. Paschall MJ, Freisthler B, Lipton RI. Moderate alcohol use and depression in young adults: findings from a national longitudinal study. Am J Public Health. 2005; 95:453–57. [PubMed: 15727976]

25. Wilsnack RW, Wilsnack SC, Kristjanson AF, Vogeltanz-Holm ND, Gmel G. Gender and alcohol consumption: patterns from the multinational GENACIS project. Addiction. 2009; 104:1487–1500. [PubMed: 19686518]

26. Dawson DA. Methodological issues in measuring alcohol use. Alcohol Res Health. 2003; 27:18–29. [PubMed: 15301397]

27. Room, R. Measuring alcohol consumption in the United States. In: Kozlowski, LT.Annis, HM.Cappell, HD.Glaser, FB.Goodstadt, MS.Israel, Y.Kalant, H.Sellers, EM., Vingilis, ER., editors. Research advances in alcohol and drug problems. Vol. 10. New York: Springer; 1990.

28. 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]

29. Little TD, Preacher KJ, Selig JP, Card Na. New developments in latent variable panel analyses of longitudinal data. Int J Behav Dev. 2007; 31:357–365.

30. Kern JL, McBride BA, Laxman DJ, Dyer WJ, Santos RM, Jeans LM. The role of multiple-group measurement invariance in family psychology research. J Fam Psychol. 2016; 30:364–74. [PubMed: 26796321]

31. Widaman, KF., Reise, SP. Exploring the measurement invariance of psychological instruments: applications in the substance use domain. In: Bryant, KJ.Windle, M., West, SG., editors. The science of prevention: methodological advances from alcohol and substance abuse research. Washington, DC: American Psychological Association; 1997.

32. Derringer J, Krueger RF, Dick DM, Agrawal A, Bucholz KK, Foroud T, et al. Measurement invariance of DSM-IV alcohol, marijuana and cocaine dependence between community-sampled and clinically overselected studies. Addiction. 2013; 108:1767–76. [PubMed: 23651171]

33. Peng C-Z, Wilsnack RW, Kristjanson AF, Benson P, Wilsnack SC. Gender differences in the factor structure of the Alcohol Use Disorders Identification Test in multinational general population surveys. Drug Alcohol Depend. 2012; 124:50–6. [PubMed: 22236536]

34. Feldstein Ewing SW, Montanaro EA, Gaume J, Caetano R, Bryan AD. Measurement invariance of alcohol instruments with Hispanic youth. Addict Behav. 2015; 46:113–20. [PubMed: 25863581]

35. Epler AJ, Sher KJ, Piasecki TM. Reasons for abstaining or limiting drinking: a developmental perspective. Psychol Addict Behav. 2009; 23:428–42. [PubMed: 19769427]

36. Cooper ML. Motivations for alcohol use among adolescents: development and validation of a four-factor model. Psychol Assess. 1994; 6:117–28.

37. Grant VV, Stewart SH, O'Connor RM, Blackwell E, Conrod PJ. Psychometric evaluation of the five-factor Modified Drinking Motives Questionnaire — Revised in undergraduates. Addict Behav. 2007; 32:2611–32. [PubMed: 17716823]

38. Gilson K-M, Bryant C, Bei B, Komiti A, Jackson H, Judd F. Validation of the Drinking Motives Questionnaire (DMQ) in older adults. Addict Behav. 2013; 38:2196–202. [PubMed: 23454884]

39. Hatzenbuehler ML, Corbin WR, Fromme K. Discrimination and alcohol-related problems among college students: a prospective examination of mediating effects. Drug Alcohol Depend. 2011; 115:213–20. [PubMed: 21145669]

40. Williams J, Jones SB, Pemberton MR, Bray RM, Brown JM, Vandermaas-Peeler R. Measurement invariance of alcohol use motivations in junior military personnel at risk for depression or anxiety. Addict Behav. 2010; 35:444–51. [PubMed: 20097012]

41. Kristjansson SD, Agrawal A, Littlefield AK, Pergadia ML, Lessov-Schlaggar CN, Sartor CE, et al. Drinking motives in female smokers: factor structure, alcohol dependence, and genetic influences. Alcohol Clin Exp Res. 2011; 35:345–54. [PubMed: 21083669]

42. Németh Z, Urbán R, Kuntsche E, Pedro EMS, Nieto JGR, Farkas J, et al. Drinking motives among Spanish and Hungarian young adults: a cross-national study. Alcohol Alcohol. 2011; 46:261–9. [PubMed: 21398656]

Author Manuscript

Author Manuscript

Author Manuscript

43. Bullers S, Cooper ML, Russell M. Social network drinking and adult alcohol involvement: a longitudinal exploration of the direction of influence. Addict Behav. 2001; 26:181–99. [PubMed: 11316376]

44. Corbin WR, Iwamoto DK, Fromme K. A comprehensive longitudinal test of the acquired preparedness model for alcohol use and related problems. J Stud Alcohol Drugs. 2011; 72:602–610. [PubMed: 21683042]

45. Nicolai J, Moshagen M, Demmel R. Patterns of alcohol expectancies and alcohol use across age and gender. Drug Alcohol Depend. 2012; 126:347–53. [PubMed: 22748519]

46. Institute of Medicine (IOM) and National Research Council (NRC). Investing in the health and well-being of young adults. Washington, DC: The National Academies Press; 2015.

47. Harris, KM. The National Longitudinal Study of Adolescent to Adult Health (Add Health), Waves I & II, 1994–1996; Wave III, 2001–2002; Wave IV, 2007–2009 [machine-readable data file and documentation]. Chapel Hill, NC: Carolina Population Center, University of North Carolina at Chapel Hill; 2009.

48. StataCorp. Stata Statistical Software: Release 14. College Station, TX: StataCorp LP; 2015.

49. Muthén, LK., Muthén, BO. Mplus user's guide Seventh Edition. Los Angeles, CA: Muthén & Muthén; 1998.

50. Chantala K. Guidelines for analyzing Add Health data. 2006

51. Cheung GW, Rensvold RB. Testing factorial invariance across groups: a reconceptualization and proposed new method. J Manag. 1999; 25:1–27.

52. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. Struct Equ Model Multidiscip J. 2002; 9:233–55.

53. Erol A, Karpyak VM. Sex and gender-related differences in alcohol use and its consequences: contemporary knowledge and future research considerations. Drug Alcohol Depend. 2015; 156:1–13. [PubMed: 26371405]

54. Byrne BM, Shavelson RJ, Muthén B. Testing for the equivalence of factor covariance and mean structures: the issue of partial measurement invariance. Psychol Bull. 1989; 105:456–66.

55. Steenkamp J-BE, Baumgartner H. Assessing measurement invariance in cross-national consumer research. J Consum Res. 1998; 25:78–107.

56. MacCallum RC, Austin JT. Applications of structural equation modeling in psychological research. Annu Rev Psychol. 2000; 51:201–226. [PubMed: 10751970]

57. Chassin, L., Colder, CR., Hussong, A., Sher, KJ. Substance use and substance use disorders. In: Cicchetti, D., editor. Developmental psychopathology. Vol. 3. New York: Wiley; 2016.

58. Zucker, RA., Hicks, BM., Heitzeg, MM. Alcohol use and the alcohol use disorders over the life course: a cross-level developmental review. In: Cicchetti, D., editor. Developmental psychopathology. New York: Wiley; 2016.

59. Hughes TL, Wilsnack SC, Kantor LW. The influence of gender and sexual orientation on alcohol use and alcohol-related problems. Alcohol Res Curr Rev. 2016; 38:121–32.

60. Delker E, Brown Q, Hasin DS. Alcohol consumption in demographic subpopulations. Alcohol Res Curr Rev. 2016; 38:7–15.

61. Collins SE. Associations between socioeconomic factors and alcohol outcomes. Alcohol Res Curr Rev. 2016; 38:83.

62. Paschall MJ, Lipperman-Kreda S, Grube JW. Effects of the local alcohol environment on adolescents' drinking behaviors and beliefs. Addiction. 2014; 109:407–16. [PubMed: 24320952]

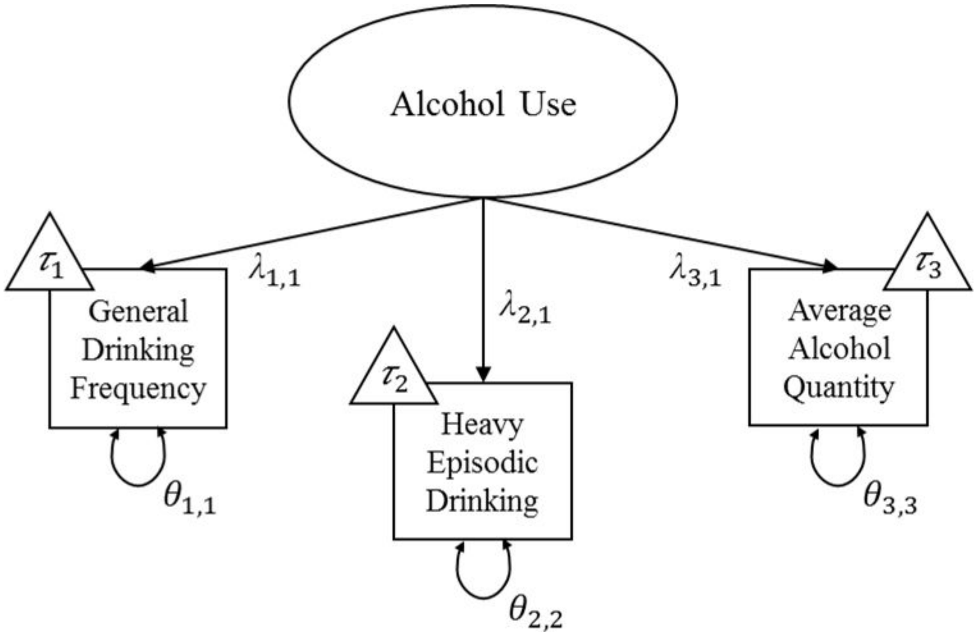

Figure 1.
Conceptual Measurement Model of an Alcohol Use Latent Construct with Three Observed Indicators.

λ = item loading; τ = item intercepts; θ = item residual variance.

Table 1

Observations, Weighted Percentages, and Mean Alcohol use by Age and Gender

	Males				Females				Total		
Age	***n***	**$\%_W$**	$\overline{M}$	**95% CI**	***n***	**$\%_W$**	$\overline{M}$	**95% CI**	***n***	$\overline{M}$	**95% CI**
13	941	48.31	.53	[.43, .62]	1,119	51.69	.45	[.37, .53]	2,060	.49	[.43, .55]
14	1,178	49.95	.77	[.64, .89]	1,333	50.05	.87	[.74, .99]	2,511	.82	[.72, .91]
15	1,603	49.98	1.19	[1.04, 1.34]	1,726	50.02	1.23	[1.10, 1.35]	3,329	1.21	[1.11, 1.31]
16	1,856	49.80	1.55	[1.41, 1.70]	1,846	50.20	1.39	[1.26, 1.53]	3,702	1.47	[1.35, 1.59]
17	1,831	52.09	2.07	[1.91, 2.23]	1,817	47.91	1.49	[1.36, 1.61]	3,648	1.79	[1.67, 1.91]
18	1,423	52.98	2.48	[2.26, 2.70]	1,395	47.02	1.61	[1.45, 1.77]	2,818	2.07	[1.92, 2.22]
19	800	50.92	2.47	[2.23, 2.71]	916	49.08	1.94	[1.73, 2.14]	1,716	2.21	[2.02, 2.39]
20	836	48.82	2.67	[2.40, 2.93]	1,061	51.18	1.81	[1.63, 1.99]	1,897	2.23	[2.03, 2.42]
21	1,026	50.46	2.64	[2.41, 2.88]	1,257	49.54	1.89	[1.72, 2.06]	2,283	2.27	[2.10, 2.44]
22	1,326	49.89	2.56	[2.35, 2.76]	1,409	50.11	1.88	[1.72, 2.04]	2,735	2.22	[2.06, 2.38]
23	1,361	51.15	2.67	[2.47, 2.88]	1,450	48.85	1.91	[1.76, 2.06]	2,811	2.30	[2.14, 2.45]
24	1,183	52.23	2.51	[2.31, 2.72]	1,225	47.77	1.68	[1.53, 1.83]	2,408	2.16	[1.97, 2.26]
25	646	53.25	2.62	[2.39, 2.85]	690	46.75	1.75	[1.55, 1.94]	1,336	2.11	[2.04, 2.38]
26	770	48.39	2.64	[2.40, 2.88]	994	51.61	1.86	[1.70, 2.03]	1,764	2.24	[2.06, 2.42]
27	949	50.68	2.56	[2.37, 2.75]	1,180	49.32	1.68	[1.56, 1.81]	2,129	2.13	[2.00, 2.26]
28	1,245	49.16	2.46	[2.29, 2.63]	1,457	50.84	1.73	[1.60, 1.86]	2,702	2.09	[1.97, 2.21]
29	1,338	50.64	2.38	[2.20, 2.56]	1,468	49.36	1.66	[1.55, 1.81]	2,806	2.02	[1.88, 2.17]
30	1,338	51.65	2.33	[2.19, 2.48]	1,446	48.35	1.51	[1.39, 1.63]	2,784	1.93	[1.83, 2.04]
31	905	54.43	2.41	[2.21, 2.61]	876	45.57	1.33	[1.19, 1.47]	1,781	1.92	[1.79, 2.05]

Note. Mean levels of alcohol use are observed averages of alcohol use frequency, HED frequency, and average drink quantity when drinking.

Table 2

Model Fit and Measurement Invariance Tests by Gender for Each Year of Age and One-Year Comparisons Among Males and Females

Measurement Invariance by Gender							Measurement Invariance by One-Year Age Increments												
									Males					Females					
Age	**Model**	χ^2	***p***	**RMSEA**	**CFI**	**ΔCFI**	**Ages**	**Model**	χ^2	***p***	**RMSEA**	**CFI**	**ΔCFI**	χ^2	***p***	**RMSEA**	**CFI**	**ΔCFI**
	C	0.00	–	.000	1.00	–		C	0.00	–	.000	1.00	–	0.00	–	.000	1.00	–
13	M	**5.50**	**.064**	**.041**	**.985**	**.015**	13 | 14	M	0.26	.878	.000	1.00	.000	4.17	.124	.030	.993	.007
	S	–	–	–	–	–		S	0.59	.965	.000	1.00	.000	8.30	.081	.030	.987	.006
14	M	2.68	.261	.017	.998	.002	14 | 15	M	0.67	.715	.000	1.00	.000	2.89	.236	.017	.999	.001
	S	7.28	.122	.026	.992	.006		S	2.23	.694	.000	1.00	.000	5.74	.219	.017	.998	.001
15	M	0.88	.644	.000	1.00	.000	15 | 16	M	0.74	.690	.000	1.00	.000	2.11	.349	.005	1.00	.000
	S	10.44	.036	.031	.995	.005		S	2.29	.683	.000	1.00	.000	3.69	.449	.000	1.00	.000
16	M	1.03	.596	.001	1.00	.000	16 | 17	M	5.13	.077	.029	.997	.003	1.19	.552	.000	1.00	.000
	S	4.02	.404	.001	1.00	.000		S	6.23	.182	.017	.998	.001	5.41	.247	.014	.999	.001
17	M	5.50	.064	.031	.997	.003	17 | 18	M	0.21	.902	.000	1.00	.000	2.68	.263	.015	.999	.001
	S	**21.11**	< **.001**	**.048**	**.985**	**.012**		S	0.65	.957	.000	1.00	.000	3.05	.549	.000	1.00	.000
18	M	3.20	.202	.021	.999	.001	18 | 19	M	1.41	.495	.000	1.00	.000	5.10	.078	.037	.998	.002
	S	**26.58**	< **.001**	**.063**	**.984**	**.015**		**S**	**14.46**	**.006**	**.049**	**.990**	**.010**	8.56	.073	.031	.997	.001
19	M	6.54	.038	.051	.995	.005	19 | 20	M	0.85	.653	.000	1.00	.000	5.20	.074	.040	.997	.003
	S	15.98	.003	.059	.988	.007		S	3.68	.451	.000	1.00	.000	12.14	.016	.045	.994	.003
20	M	8.74	.013	.060	.994	.006	20 | 21	M	1.83	.401	.000	1.00	.000	2.96	.228	.020	.999	.001
	S	**27.69**	< **.001**	**.079**	**.978**	**.016**		S	11.67	.020	.045	.994	.006	14.15	.007	.047	.990	.009
21	M	4.51	.105	.033	.998	.002	21 | 22	M	5.67	.059	.039	.998	.002	0.37	.832	.000	1.00	.000
	S	15.89	.003	.051	.989	.009		S	16.29	.003	.051	.992	.006	4.96	.291	.013	.999	.001
22	M	6.48	.039	.040	.996	.004	22 | 23	M	0.41	.816	.000	1.00	.000	3.31	.191	.021	.998	.002
	S	16.80	.002	.048	.988	.008		S	1.24	.871	.000	1.00	.000	3.47	.483	.000	1.00	.000

Measurement Invariance by Gender							Measurement Invariance by One-Year Age Increments												
									Males					Females					
Age	**Model**	χ^2	***p***	**RMSEA**	**CFI**	**ΔCFI**	**Ages**	**Model**	χ^2	***p***	**RMSEA**	**CFI**	**ΔCFI**	χ^2	***p***	**RMSEA**	**CFI**	**ΔCFI**
23	M	**16.60**	< **.001**	**.072**	**.988**	**.012**	23 | 24	M	0.61	.737	.000	1.00	.000	0.41	.814	.000	1.00	.000
	S	–	–	–	–	–		S	1.52	.823	.000	1.00	.000	1.97	.741	.000	1.00	.000
24	M	**17.76**	< **.001**	**.081**	**.988**	**.012**	24 | 25	M	2.97	.227	.023	.999	.001	5.65	.059	.044	.996	.004
	S	–	–	–	–	–		S	6.88	.143	.028	.998	.001	**18.68**	**.001**	**.062**	**.983**	**.013**
25	M	**12.52**	**.002**	**.089**	**.986**	**.014**	25 | 26	M	0.69	.710	.000	1.00	.000	2.99	.224	.024	.998	.002
	S	–	–	–	–	–		S	1.91	.752	.000	1.00	.000	4.22	.378	.008	1.00	.000
26	M	**11.75**	**.003**	**.074**	**.986**	**.014**	26 | 27	M	0.04	.978	.000	1.00	.000	5.66	.059	.041	.995	.005
	S	–	–	–	–	–		S	0.16	.997	.000	1.00	.000	6.61	.158	.024	.996	.004
27	M	**16.95**	< **.001**	**.084**	**.984**	**.016**	27 | 28	M	2.21	.331	.010	1.00	.000	7.94	.019	.047	.993	.007
	S	–	–	–	–	–		S	4.01	.405	.001	1.00	.000	10.66	.031	.036	.992	.001
28	M	1.89	.390	.000	1.00	.000	28 | 29	M	0.78	.678	.000	1.00	.000	3.46	.178	.022	.998	.002
	S	10.80	.029	.035	.992	.008		S	2.57	.633	.000	1.00	.000	4.25	.374	.006	1.00	.000
29	M	5.95	.051	.038	.997	.003	29 | 30	M	0.32	.852	.000	1.00	.000	2.88	.237	.017	.999	.001
	S	13.81	.008	.042	.992	.005		S	2.51	.643	.000	1.00	.000	6.25	.182	.020	.996	.004
30	M	**14.69**	**.001**	**.068**	**.981**	**.019**	30 | 31	M	0.34	.842	.000	1.00	.000	**7.02**	**.030**	**.046**	**.989**	**.011**
	S	–	–	–	–	–		S	0.68	.954	.000	1.00	.000	–	–	–	–	–
31	M	**12.51**	**.002**	**.077**	**.984**	**.016**												
	S	–	–	–	–	–												

C = Configural; M = Metric; S = Scalar. All models passed configural invariance. Bolded values indicate non-invariance at ΔCFI < .010. ΔCFI for metric models are in comparison to the configural CFI and ΔCFI for scalar models are in comparison to the metric CFI. Overall chi-square test of model fit is *df* = 2 for metric and *df* = 4 for scalar models. | indicates a comparison (e.g., 13 | 14 compares measurement invariance between 13 and 14 year olds).

Table 3

Loadings, Intercepts, and Residual Variances for Males and Females by Age and Item Level Measurement Invariance Tests for Males and Females in One-Year Increments and between Males and Female Within Each Year of Age.

		Male One-Year Comparisons					Female One-Year Comparisons					Male v. Female	
					Metric	Scalar				Metric	Scalar	Metric	Scalar
Age	Item	λ	τ	θ	ΔCFI	ΔCFI	λ	τ	θ	ΔCFI	ΔCFI	ΔCFI	ΔCFI
13	DF	.98	.49	.33	✓	✓	.90	.49	.25	✓	✓	**.000**	✗
	HD	.80	.26	.28	✓	✓	.54	.19	.25	✓	✓	.011	✗
	AQ	1.97	.84	.84	✓	✓	1.44	.70	.70	✓	✓	.013	✗
14	DF	1.05	.69	.34	✓	✓	1.04	.79	.45	✓	✓	✓	✓
	HD	.91	.40	.39	✓	✓	.80	.39	.35	✓	✓	✓	✓
	AQ	2.05	1.22	1.82	✓	✓	2.08	1.42	2.22	✓	✓	✓	✓
15	DF	1.35	1.00	.34	✓	✓	1.21	1.10	.42	✓	✓	✓	✓
	HD	1.15	.65	.54	✓	✓	1.01	.57	.42	✓	✓	✓	✓
	AQ	2.43	1.94	3.70	✓	✓	2.29	2.01	2.64	✓	✓	✓	✓
16	DF	1.43	1.28	.46	✓	✓	1.29	1.21	.42	✓	✓	✓	✓
	HD	1.30	.86	.53	✓	✓	1.12	.72	.48	✓	✓	✓	✓
	AQ	2.73	2.56	4.33	✓	✓	2.28	2.26	3.67	✓	✓	✓	✓
17	DF	1.52	1.66	.55	✓	✓	1.24	1.36	.66	✓	✓	✓	**.000**
	HD	1.48	1.20	.48	✓	✓	1.10	.74	.43	✓	✓	✓	**.008**
	AQ	2.80	3.37	5.18	✓	✓	2.35	2.37	2.54	✓	✓	✓	.010
18	DF	1.64	1.97	.54	✓	**.003**	1.30	1.47	.61	✓	✓	✓	**.001**
	HD	1.62	1.51	.53	✓	.012	1.16	.83	.46	✓	✓	✓	.011
	AQ	3.00	4.02	5.78	✓	**.002**	2.26	2.54	3.34	✓	✓	✓	.011
19	DF	1.65	2.13	.77	✓	✓	1.47	1.82	.45	✓	✓	✓	✓
	HD	1.55	1.56	.49	✓	✓	1.30	1.07	.52	✓	✓	✓	✓
	AQ	2.87	3.74	4.39	✓	✓	2.19	2.92	3.21	✓	✓	✓	✓
20	DF	1.77	2.34	.58	✓	✓	1.43	1.77	.41	✓	✓	✓	**.000**

		Male One-Year Comparisons					Female One-Year Comparisons					Male v. Female	
					Metric	Scalar				Metric	Scalar	Metric	Scalar
Age	**Item**	λ	τ	θ	**ΔCFI**	**ΔCFI**	λ	τ	θ	**ΔCFI**	**ΔCFI**	**ΔCFI**	**ΔCFI**
	HD	1.62	1.64	.57	✓	✓	1.10	.91	.58	✓	✓	✓	**.007**
	AQ	2.91	4.04	4.62	✓	✓	1.96	2.76	3.65	✓	✓	✓	.014
21	DF	1.62	2.48	.91	✓	✓	1.39	2.01	.90	✓	✓	✓	✓
	HD	1.58	1.65	.59	✓	✓	1.22	1.01	.53	✓	✓	✓	✓
	AQ	2.66	3.82	4.87	✓	✓	1.97	2.66	3.14	✓	✓	✓	✓
22	DF	1.56	2.58	1.05	✓	✓	1.29	2.08	1.04	✓	✓	✓	✓
	HD	1.58	1.62	.47	✓	✓	1.10	.93	.47	✓	✓	✓	✓
	AQ	2.30	3.49	4.13	✓	✓	1.88	2.63	2.95	✓	✓	✓	✓
23	DF	1.58	2.66	1.01	✓	✓	1.42	2.08	.64	✓	✓	**.002**	✘
	HD	1.63	1.75	.40	✓	✓	1.09	.96	.69	✓	✓	**.002**	✘
	AQ	2.28	3.60	4.56	✓	✓	1.79	2.70	3.56	✓	✓	.011	✘
24	DF	1.56	2.56	1.03	✓	✓	1.27	1.90	.80	✓	.011	**.006**	✘
	HD	1.56	1.59	.55	✓	✓	.96	.77	.53	✓	**.001**	**.000**	✘
	AQ	2.16	3.41	4.26	✓	✓	1.66	2.36	3.06	✓	**.009**	.012	✘
25	DF	1.43	2.64	1.05	✓	✓	1.48	1.96	.62	✓	✓	**.002**	✘
	HD	1.40	1.56	.67	✓	✓	1.18	.99	.61	✓	✓	.013	✘
	AQ	2.22	3.67	4.20	✓	✓	1.63	2.29	2.59	✓	✓	**.005**	✘
26	DF	1.54	2.73	1.19	✓	✓	1.34	2.14	.85	✓	✓	**.006**	✘
	HD	1.55	1.61	.58	✓	✓	1.23	1.09	.43	✓	✓	.011	✘
	AQ	2.28	3.60	4.48	✓	✓	1.51	2.37	2.58	✓	✓	**.002**	✘
27	DF	1.53	2.66	1.22	✓	✓	1.41	1.99	1.03	✓	✓	**.000**	✘
	HD	1.55	1.52	.54	✓	✓	1.10	.93	.51	✓	✓	.012	✘
	AQ	2.24	3.51	4.75	✓	✓	1.58	2.13	2.38	✓	✓	.014	✘
28	DF	1.54	2.65	1.15	✓	✓	1.29	1.98	1.16	✓	✓	✓	✓
	HD	1.45	1.46	.61	✓	✓	1.25	.95	.35	✓	✓	✓	✓

Age	Item	Male One-Year Comparisons λ	τ	θ	Metric ΔCFI	Scalar ΔCFI	Female One-Year Comparisons λ	τ	θ	Metric ΔCFI	Scalar ΔCFI	Male v. Female Metric ΔCFI	Scalar ΔCFI
	AQ	2.02	3.27	4.79	✓	✓	1.54	2.26	2.97	✓	✓	✓	✓
29	DF	1.54	2.55	1.15	✓	✓	1.38	1.91	.91	✓	✓	✓	✓
	HD	1.51	1.45	.52	✓	✓	1.16	.91	.43	✓	✓	✓	✓
	AQ	1.97	3.16	4.40	✓	✓	1.50	2.15	2.85	✓	✓	✓	✓
30	DF	1.52	2.55	1.22	✓	✓	1.45	1.83	.69	.012	✘	**.000**	✘
	HD	1.54	1.45	.44	✓	✓	1.10	.81	.57	**.005**	✘	.014	✘
	AQ	2.00	3.01	3.36	✓	✓	1.29	1.89	2.25	**.000**	✘	.031	✘
31	DF	1.60	2.59	1.21	✓	✓	1.34	1.68	1.06	✓	✓	**.005**	✘
	HD	1.57	1.51	.56	✓	✓	.97	.66	.42	✓	✓	**.000**	✘
	AQ	2.05	3.14	4.18	✓	✓	1.54	1.65	1.77	✓	✓	.019	✘

Note. All models passed configural invariance. Bold values denote an improvement in model fit when item is estimated at ΔCFI < .010. ΔCFI values reflect comparisons between that age year and the age directly following that age year.

✓ = omnibus tests reflect metric or scalar invariance, thus item loadings and intercepts are invariant across comparisons

✘= when metric invariance failed, scalar variance was not tested.

DF = drinking frequency; HD = HED frequency; AQ = average quantity when drinking.

λ = item loading; τ = item intercepts; θ = item residual variance.

Table 4

Model Fit and Measurement Invariance Tests by Age in Two-Year Increments for Males and Females

		Males					Females				
Ages	**Model**	χ^2	***p***	**RMSEA**	**CFI**	**ΔCFI**	χ^2	***p***	**RMSEA**	**CFI**	**ΔCFI**
	C	0.00	–	.000	1.00	–	0.00	–	.000	1.00	–
13–14 | 15–16	M	2.08	.354	.004	1.00	.000	10.42	.006	.037	.993	.007
	S	9.48	.050	.022	.996	.004	19.12	.007	.035	.988	.005
15–16 | 17–18	M	11.93	.003	.038	.996	.004	0.99	.634	.000	1.00	.000
	S	20.05	.001	.035	.994	.002	5.78	.216	.011	.999	.001
17– 18 | 19– 20	M	5.65	.059	.027	.999	.001	16.34	< .000	.053	.995	.005
	S	**35.74**	< **.001**	**.057**	**.987**	**.012**	32.45	< .001	.052	.989	.006
19– 20 | 21– 22	M	11.94	.003	.050	.997	.003	0.78	.679	.000	1.00	.000
	S	**63.55**	< **.001**	**.086**	**.980**	**.017**	**45.57**	< **.001**	**.067**	**.982**	**.018**
21– 22 | 23– 24	M	8.03	.018	.035	.998	.002	5.28	.071	.025	.998	.002
	S	14.61	.006	.033	.997	.001	8.08	.089	.020	.998	.000
23– 24 | 25– 26	M	2.57	.277	.012	1.00	.000	13.15	.001	.051	.992	.008
	S	13.43	.009	.035	.997	.003	**39.04**	< **.001**	**.063**	**.975**	**.017**
25– 26 | 27– 28	M	2.04	.361	.003	1.00	.000	0.46	.796	.000	1.00	.000
	S	4.54	.338	.009	1.00	.000	1.12	.892	.000	1.00	.000
27– 28 | 29– 30	M	3.01	.223	.014	1.00	.000	4.53	.104	.021	.998	.002
	S	10.95	.027	.027	.997	.003	6.81	.146	.016	.998	.000
29– 30 | 31	M	0.13	.938	.000	1.00	.000	7.21	.027	.037	.994	.006
	S	0.20	.996	.000	1.00	.000	11.33	.023	.031	.992	.002

Notes. All models passed configural invariance. Bolded values indicate non-invariance at ΔCFI < .010. C = Configural; M = Metric; S = Scalar. Model chi-square test of model fit is *df* = 2 for metric and *df* = 4 for scalar models. | indicates a comparison: For example, 13–14 | 15–16 compares measurement invariance of 13 and 14-year-olds to 15 and 16-year-olds.

Table 5

Loadings, Intercepts, and Residual Variances for Males and Females by Age and Item Level Measurement Invariance Tests for Males and Females in Two-Year Increments

		Male Two-Year Comparisons					Female Two-Year Comparisons				
					Metric	Scalar				Metric	Scalar
Ages	Item	λ	τ	θ	ΔCFI	ΔCFI	λ	τ	θ	ΔCFI	ΔCFI
13–14	DF	1.02	.60	.33	✓	✓	.98	.65	.37	✓	✓
	HD	.86	.33	.34	✓	✓	.69	.29	.30	✓	✓
	AQ	2.02	1.05	1.37	✓	✓	1.84	1.07	1.49	✓	✓
15–16	DF	1.40	1.14	.40	✓	✓	1.25	1.16	.42	✓	✓
	HD	1.23	.75	.53	✓	✓	1.07	.65	.45	✓	✓
	AQ	2.60	2.25	4.02	✓	✓	2.29	2.13	3.16	✓	✓
17–18	DF	1.59	1.81	.55	✓	**.002**	1.27	1.41	.64	✓	✓
	HD	1.56	1.35	.50	✓	.011	1.13	.78	.45	✓	✓
	AQ	2.91	3.68	5.47	✓	**.006**	2.30	2.45	2.92	✓	✓
19–20	DF	1.72	2.24	.67	✓	**.003**	1.44	1.79	.44	✓	**.000**
	HD	1.58	1.61	.53	✓	.016	1.20	.98	.56	✓	.012
	AQ	2.90	3.90	4.52	✓	**.001**	2.07	2.83	3.45	✓	.010
21–22	DF	1.59	2.53	.99	✓	✓	1.34	2.04	.98	✓	✓
	HD	1.59	1.63	.53	✓	✓	1.16	.97	.50	✓	✓
	AQ	2.48	3.66	4.56	✓	✓	1.93	2.65	3.04	✓	✓
23–24	DF	1.57	2.61	1.02	✓	✓	1.34	1.99	.72	✓	.017
	HD	1.59	1.67	.48	✓	✓	1.03	.86	.61	✓	**.005**
	AQ	2.22	3.50	4.41	✓	✓	1.73	2.53	3.31	✓	**.005**
25–26	DF	1.49	2.69	1.13	✓	✓	1.40	2.07	.77	✓	✓
	HD	1.49	1.59	.62	✓	✓	1.21	1.05	.51	✓	✓
	AQ	2.25	3.63	4.37	✓	✓	1.56	2.34	2.58	✓	✓
27–28	DF	1.53	2.65	1.19	✓	✓	1.35	1.99	1.10	✓	✓

		Male Two-Year Comparisons					Female Two-Year Comparisons				
					Metric	Scalar				Metric	Scalar
Ages	**Item**	**λ**	**τ**	**θ**	**ΔCFI**	**ΔCFI**	**λ**	**τ**	**θ**	**ΔCFI**	**ΔCFI**
	HD	1.50	1.49	.57	✓	✓	1.17	.94	.44	✓	✓
	AQ	2.13	3.39	4.78	✓	✓	1.56	2.20	2.69	✓	✓
29–30	DF	1.53	2.55	1.19	✓	✓	1.41	1.87	.81	✓	✓
	HD	1.53	1.45	.48	✓	✓	1.13	.86	.50	✓	✓
	AQ	2.00	3.08	3.87	✓	✓	1.40	2.02	2.57	✓	✓
31	DF	1.60	2.59	1.21	✓	✓	1.34	1.68	1.06	✓	✓
	HD	1.57	1.51	.56	✓	✓	.97	.66	.42	✓	✓
	AQ	2.05	3.14	4.18	✓	✓	1.54	1.65	1.77	✓	✓

Note. Bold values denote an improvement in model fit when item is estimated at ΔCFI < .010.

ΔCFI values reflect comparisons between that age year and the age directly following that age year.

✓ = omnibus tests reflect metric or scalar invariance, thus item loadings and intercepts are invariant across comparisons

DF = drinking frequency; HD = HED frequency; AQ = average quantity when drinking.

λ = item loading; τ = item intercepts; θ = item residual variance.