Published in final edited form as:

Alcohol Clin Exp Res. 2009 April ; 33(4): 663–673. doi:10.1111/j.1530-0277.2008.00881.x.

A Rasch Model Analysis of Alcohol Consumption and Problems Across Adolescence and Young Adulthood

Christopher W. Kahler, Bettina B. Hoepfner, and Kristina M. Jackson

Center for Alcohol and Addiction Studies, Brown University, Providence, Rhode Island.

Abstract

Background—Recent investigations using item response modeling have begun to conceptualize alcohol consumption, problems, and dependence as representing points along a single continuum of alcohol involvement. Such a conceptualization may be of particular benefit to measurement of alcohol involvement in adolescents, but investigations to date have been limited to adult samples and may not generalize to adolescents due to age-related developmental differences.

Methods—This study used Rasch model analyses to examine the properties of indices of alcohol consumption and problems among 6,353 adolescents, aged 12 to 18 years, in Wave 1 of the Add Health survey. A particular focus was on whether the functioning of items changed when these adolescents were re-interviewed in Wave 3 when they were 18 to 24 years of age.

Results—Rasch model analyses supported the unidimensionality and additive properties of the items in the Wave 1 data. Comparisons of Wave 1 and Wave 3 data indicated differential item functioning in most of the items such that items related to alcohol consumption were more severe during adolescence, whereas items related to alcohol problems were more severe in young adulthood.

Conclusions—A valid index of alcohol involvement in adolescents can be constructed combining indices of alcohol consumption and alcohol problems. Such an index covers a range of severity and functions similarly across sex and race/ethnicity. A similar index can be constructed in young adulthood. However, the interpretation of scores must be attentive to developmental differences. In particular, for adolescents, indices of alcohol consumption are relatively closer in severity to indices of alcohol problems than they are among young adults. Thus, alcohol problems are more likely among adolescents than young adults given a similar level of drinking.

Keywords

Alcohol Consumption; Alcohol Problems; Rasch Modeling; Adolescence; Measurement

ALCOHOL CONSUMPTION, ALCOHOL-RELATED problems, and alcohol use disorders have typically been measured and analyzed as separate phenomena. Such an approach can be justified given the clear conceptual distinctions among constructs such as intensity of alcohol use, alcohol-related problems, loss of control over drinking, and physiological dependence. At the same time, there also may be substantial benefits to creating continuous measures of alcohol involvement that incorporate multiple dimensions of alcohol use, problems, and dependence in a single index. Such an index provides greater information across a full range of alcohol involvement compared to relying on diagnostic criteria alone (Krueger et al., 2004) and allows for more comprehensive and unified examination of predictors and correlates of alcohol

involvement in community samples (Kahler et al., 2008). Recent studies using item response modeling (IRM) have demonstrated that it is psychometrically justifiable to scale measures of (1) drinking quantity and frequency and (2) symptoms of alcohol abuse and dependence along a single latent dimension (Kahler et al., 2008; Krueger et al., 2004; Saha et al., 2007). Based on this evidence, 2 recent papers have argued that future diagnostic systems should incorporate measures of harmful patterns of drinking into criteria for the alcohol dependence syndrome (Li et al., 2007; Saha et al., 2007).

Although a combined index of regular excessive drinking and alcohol use disorder criteria may provide reliable information about clinically relevant alcohol involvement in adult samples, such a measure may be of less utility in studies of adolescents who are early in their drinking careers. In particular, many alcohol use disorder symptoms (e.g., withdrawal) may be rare in adolescents that they provide limited information about the vast majority of those adolescents who have consumed alcohol. For example, Neal and colleagues (2006) found that the majority of items on the Rutgers Alcohol Problem Index (White and Labouvie, 1989) assessed alcohol problems that were so severe that the measure provided relatively little information about adolescents in the lower and middle regions of the alcohol involvement continuum. In community adolescent samples, creating a single alcohol involvement index by combining measures of early milestones of alcohol consumption (e.g., drinking to intoxication or drinking at least 12 drinks in a year) with measures of early milestones of intoxication/problem drinking (e.g., having a hangover or a regretted behavior while drinking) may provide a more sensitive assessment of the initial trajectories of alcohol involvement in adolescents. Results of item response analyses have indicated that mild alcohol problems and symptoms of acute intoxication do index a common latent continuum that is also indexed by symptoms of alcohol dependence (Kahler et al., 2003a,b, 2005; Krueger et al., 2004) with items differing quantitatively rather than qualitatively. However, these results have been obtained using only young adult and general adult samples. It is not clear to which extent the results obtained in adults will generalize to adolescents.

Methods used to assess alcohol involvement in adolescents may perform differently than they do in young adults and older adults due to important developmental differences (Chung et al., 2005; Martin et al., 2006). For example, there is a significant increase in alcohol involvement from high school to college (Baer et al., 1995; Sher and Rutledge, 2007), a time during which certain indices of alcohol involvement become increasingly normative while others remain relatively rare and indicative of significant pathology. Consequently, items tapping relatively low levels of alcohol consumption (e.g., drinking at least once a month) may indicate relatively higher levels of severity of alcohol involvement in adolescence compared to young adulthood, and therefore have different prognostic significance regarding whether indices of more pathological alcohol involvement are likely to be endorsed. Drift in item severity (i.e., the level of severity at which an item has a 50% likelihood of being endorsed) across ages can occur even if the items continue to show consistent loadings on a single latent factor. For example, using IRM, Neal and colleagues (2006) found that both missing school and failing to do homework assignments were indicative of relatively more severe alcohol problems in high school than they were in college, whereas getting into arguments with parents was a less severe symptom in high school than it was in college.

Differential item functioning (DIF) by age (which serves as a proxy for developmental stage) and school setting has important conceptual implications as it suggests which indicators of alcohol involvement are most sensitive to developmental and environmental contexts. Also, if there is substantial drift in the relative severity or ordering of items with age, the substantive interpretation of a given total score may vary with age given that the items that typically lead to the score will also vary. For example, a given score at one age may reflect likely endorsement only of regular and heavy drinking, whereas at another age, the same score may indicate a

substantially higher odds of endorsing negative drinking consequences. Conversely, even as overall rates of endorsement of alcohol items increase across adolescence and into young adulthood, certain items may maintain their relative ordering and their relationship to the latent continuum quite well such that the same items would be expected to be endorsed with a given total score regardless of the age of the respondent; such items serve as good anchors when examining alcohol involvement over time and across age groups.

Study Aims

The purpose of the present study was to conduct item response analyses using the Rasch model (Rasch, 1960) to examine (1) whether a psychometrically sound additive index of alcohol involvement can be created by combining indices of alcohol consumption and problems among adolescents and (2) how the performance of items in this index differs across demographic groups and across early adolescence, middle adolescence, and young adulthood. Data were drawn from the National Longitudinal Study of Adolescent Health (Add Health; Udry, 2003), a longitudinal study designed to examine social-environmental influences on adolescent health, including behaviors such as substance use. Although Add Health lacks measures of alcohol abuse and dependence, it provides a large longitudinal community sample of adolescents and contains items related to alcohol consumption and alcohol problems that are relevant for mapping the alcohol involvement continuum in this age group.

We first examined the characteristics of 15 indices of alcohol consumption and problems among all participants (aged 12 to 18 years) in the initial wave of the study (Wave 1) and explored whether items performed similarly across sex and ethnic groups. Given the developmental framework of the study, we then considered the extent to which age altered item functioning by comparing item parameters obtained from the sample at the baseline interview with parameters obtained from the sample at a 6-year follow-up (Wave 3) when participants were no longer in high school and were of the ages of 18 to 24 years. We also conducted comparisons within each wave by comparing 12- to 14-year-olds with 15- to 18-year-olds in Wave 1, and comparing 18- to 20-year-olds with 21- to 24-year-olds in Wave 3. We expected to find greater differences in item functioning across Waves 1 and 3 than we would when making age comparisons within each wave; the Wave 1 to Wave 3 comparisons involve greater differences in age and more developmentally significant differences in environmental contexts given that participants transitioned out of high school between these waves. In particular, we expected that items tapping light or moderate levels of alcohol consumption would indicate relatively lower levels of severity of alcohol involvement in young adulthood than they do during adolescence. We further expected these differences in severity thresholds for certain items would emerge even when we controlled for differences in alcohol exposure across ages by limiting analyses to participants who had had at least some significant alcohol use in the past year.

MATERIALS AND METHODS

Data Collection

The National Longitudinal Study of Adolescent Health (Add Health) is a nationally representative study funded by the National Institute of Child Health and Human Development and other federal agencies that explore the causes of health-related behaviors of adolescents in grades 7 through 12 and their outcomes in young adulthood. Data were collected in 3 waves, starting in 1994 with an in-school administered questionnaire and in-home interviews (Wave 1), and continuing in 1996 (Wave 2) and 2001 to 2002 (Wave 3) with at-home administered interviews (see <http://www.cpc.unc.edu/projects/addhealth> for additional details). The Add Health survey used a multistage cluster design, stratified by region, urbanicity, and school type, ethnic mix, and size. The final sample includes students from private, religious, and public

schools located in urban, suburban, and rural areas. All students present on the day of assessment participated in a 45-minute long in-school survey. Following the initial (Wave 1) survey, the study collected follow-up information via in-home interviews. The present study drew on Waves 1 and 3 (6-year follow-up) of the publicly available Add Health dataset which includes a random selection of the core sample and an oversample of African-American adolescents with a parent who has a college degree. Add Health includes sample weights to account for the complex sampling design. These weights are not used in the current analyses, as our purpose was not to make generalizations about the U.S. population but rather to examine the relationships of individual items to an underlying continuum and whether these differed across subgroups.¹

Given our focus on age differences and development, we restricted our analyses in 2 additional ways. First, we excluded participants from analyses who were unusually young (i.e., less than 12 years of age, $n = 4$) or unusually old (i.e., older than 18 years of age, $n = 144$) at Wave 1, so as to define our age ranges more precisely. Second, we excluded participants who were still in high school ($n = 11$) at Wave 3, so that Wave 3 responses did not reflect high school experiences. The restricted Wave 1 sample consisted of $n = 6,353$ adolescents and young adults (51.8% female; 59.7% non-Hispanic White; 23.7% African American/Black; 11.3% Hispanic/Latino; 3.4% Asian American; 0.9% American Indian; 0.9% of other ethnicity) between 12 and 18 years of age at Wave 1. Of these, 4,761 were reinterviewed as young adults between 18 and 24 years of age at Wave 3. Wave 3 completion rates were significantly higher in females compared to males (78.6% vs. 71.8%) and in the younger age group (12 to 14 years) compared to the older age group (15 to 18 years; 80.7% vs. 72.8%). Whites had significantly higher completion rates than Hispanic/Latinos (76.5% vs. 69.2%) but did not differ from African Americans (74.8%) or other race/ethnicity groups (76.7%). The nonparametric Kruskal–Wallis test indicated that Wave 3 completers endorsed significantly fewer alcohol involvement items at Wave 1 (median = 0, interquartile range = 0–3) compared to Wave 3 noncompleters (median = 0, interquartile range = 0–4), $p = 0.007$. Alcohol involvement did not significantly predict Wave 1 completion status, however, when controlling for demographic variables.

Measures

Sex, race, and age were assessed at Wave 1. From age, we created several age groups within adolescence and within young adulthood. The adolescent age group distinction was between younger (age 12 to 14 years) and older (age 15 to 18 years) adolescents, splitting the sample at age 14 given that the average age of drinking onset is 14 years old (Grunbaum et al., 2004). The age groups within emerging adulthood corresponded to those who had not yet reached legal drinking age (young emergent adults; 18 to 20 years of age) versus those who had attained legal drinking status (older emergent adults; 21 to 24 years of age).

The Wave 1 Add Health survey included a section assessing alcohol, cigarette, and illicit drug use. Alcohol use was assessed with 17 items; the present study used 13 items which assessed experiences within the past 12 months, as opposed to lifetime experiences (a 14th item was omitted because it queried the type of drinks consumed, and thus is not a milestone of alcohol involvement). We excluded the 3 items that assessed lifetime experiences as these would by necessity be cumulative over time.

From the remaining 13 items, 15 dichotomous items were created that reflect milestone indicators of alcohol involvement. Items were dichotomized for 2 reasons. First, for all items

¹To determine whether conducting weighted analyses would meaningfully alter results, we compared the unweighted Rasch model results at Wave 1 with weighted analyses in which weights were adjusted so that sample size reflected the total number of subjects in the analysis rather than the population. The absolute value of differences in severity estimates across the 2 analyses for the 15 items analyzed averaged only 0.04 logits, range = 0.0 to 0.09, which represents minimal differences.

the most commonly endorsed response was “never” or “none,” and all items showed a positively skewed triangular distribution in which very few participants endorsed the highest response option of either “every day or almost every day” or “5 or more times” (e.g., <1% of the sample endorsed the highest option for all problem variables except hung-over and threw up). Second, consistent with other recent studies involving scaled items (Kahler et al., 2004; Neal et al., 2006), we preferred the simplicity of coding items as absent or present in the past 12 months as this is consistent with viewing the items as milestone markers of alcohol involvement that can be easily interpreted when summed. Eleven of these 15 milestone indicators were coded from unique questions. Three binary variables were derived from the item “During the past 12 months, on how many days did you drink alcohol?”: daily drinking (if participants responded “every day or almost every day”), weekly drinking (if participants responded “1 or 2 days a week” or more), and monthly drinking (if participants responded “2 or 3 days a month” or more). A question of typical drinking quantity was used in conjunction with the drinking frequency item to create a “12 or more drinks” in a year item, which has been used as an indicator of minimal alcohol involvement in a prior study (Kahler et al., 2008).

Wave 3 items were highly similar but not identical to those assessed in Wave 1. Two items (“parent problem” and “regretted action”) were dropped, and 1 item “drunk at school or work” was added in the Add Health survey. Wave 3 items were coded in the same fashion as Wave 1 items, resulting in 14 binary milestone markers of alcohol involvement.

Data Analysis Plan

We conducted Rasch model analyses of the 15 alcohol involvement indicators using the BIGSTEPS software package (Linacre and Wright, 1998). The Rasch model is a 1-parameter logistic item response model that scales both items and persons along a single latent dimension using an equal interval log-odds scale. The odds of an individual endorsing a given item is modeled as a function of the individual's overall level of problem severity and the severity of that item (Wright and Masters, 1982). In this model, item severity is defined by the point on a latent severity continuum at which the item has a 50% likelihood of being endorsed. The Rasch model differs from 2-parameter logistic (2-PL) item response model in that it does not include estimation of a discrimination parameter (i.e., the ability of the item to discriminate between individuals above vs. below the item's severity threshold). Rather than model item responses with extra parameters, the Rasch model uses fit statistics to determine which items fit the specification that discrimination parameters across the items are essentially equal. The advantage to this approach is that with fit to the Rasch model, the ordering of persons and items can be assumed to hold such that (1) item A is a more severe (i.e., less likely to be endorsed) symptom than item B regardless of the severity of the respondent and (2) an individual with greater severity is more likely to endorse any given test item than a person with lower severity. Furthermore, with fit to the model, an unweighted sum of items can serve as a sufficient measure of an individual's severity, which is consistent with how items are typically used in scales.

We examined fit statistics for each item to determine how well the data fit the Rasch model (Wright and Masters, 1982). Values ranging from 0.60 to 1.40 are considered to indicate adequate model fit for survey rating scales such as this one (Linacre and Wright, 1994); 1.0 represent ideal fit for the Rasch model, less than 1.0 represent less variability (randomness) than expected, and greater than 1.0 indicates more variability than expected. Infit values are weighted so as to give more relative weight to the performances of persons close to the item's severity. They are therefore less susceptible to outlier influences than outfit statistics and generally preferred (Bond and Fox, 2001). Fit statistics provide information that is very highly correlated with discrimination parameters from 2-PL models in the reverse direction (i.e., lower infit is related to higher discrimination) (Kahler and Strong, 2006).² We also examined the

point biserial correlations between a given item and the total of the remaining items; this serves as an alternative index of how strongly an item relates to levels of alcohol involvement.

To test the Rasch model assumption that responses to a given item are independent from responses to other items (i.e., locally independent), we examined the residuals for each item after removing the contribution of a single latent dimension to item responses. Local independence is satisfied if item residuals show low correlations with each other (Wright, 1996). We also conducted a principal components analysis of these residuals. Factors from residual analysis accounting for >1.5 units of variance (i.e., eigenvalues) are considered significant (Linacre, 1998; Smith and Miao, 1994), and the absence of significant factors in the residuals supports the assumption of local independence.

After conducting Rasch model analyses in the full sample of 12 to 18 year olds, we examined DIF across demographic subgroups to determine whether the alcohol involvement items showed any bias by group. The estimation of DIF involves comparing analyses conducted separately within each demographic group (Holland and Wainer, 1993). If items behave similarly across groups, then severity parameters estimated independently in different samples will fall within an acceptable range of agreement (e.g., 95% confidence interval), and the significance of the difference between the estimates can be tested with a z-test using the pooled standard error of the estimates. We compared (1) girls to boys, (2) non-Hispanic White to Black participants, (3) non-Hispanic White to Hispanic/Latino participants, and (4) Black to Hispanic participants. Comparisons were not conducted with Asian, American Indian, and other ethnicities as the numbers of these participants endorsing at least one alcohol involvement indicator were too low for reliable estimation ($n = 65, 23, \text{ and } 20$, respectively).

For each DIF test, we controlled the family-wise error rate using the Hochberg (Hochberg, 1988) procedure. This procedure assumes that p -values are independent and uniformly distributed under their respective null hypotheses, and contrasts the ordered p -values with the same set of critical values. It rejects all hypotheses with smaller or equal p -values to that of any one found less than its critical value, and thereby controls the family-wise error rate in *step-up* fashion, as opposed to similar but less powerful step-down procedures (Holm, 1979). This adjustment to p -values helps maintain desired family-wise error rates when conducting multiple tests of significance and is a more statistically powerful alternative to the overly conservative Bonferroni adjustment. Given the relatively large sample sizes involved in the present study, differences in severity estimates between groups could be significant statistically but not be of sufficient magnitude to affect the interpretation of scores across groups or to be conceptually meaningful. We therefore relied on a common convention for Rasch model analysis by considering only items with significant DIF greater than 0.50 logits to be meaningful as items with DIF of 0.50 logits or less are unlikely to influence the overall test (Draba, 1977; Lai et al., 2005; Tristan, 2006). We labeled a specific item as functioning significantly differentially only if (1) the DIF t -test was significant at the 0.05 (2-tailed) level after adjusting for multiple comparisons using the Hochberg (1988) procedure, and (2) if DIF exceed 0.50 logits.

After we examined DIF with Wave 1, we conducted a Rasch model analysis of Wave 3 data. Analyses were conducted using all available data at that wave rather than using only those subjects who were also in the Wave 1 Rasch model analyses. We then conducted analyses to test DIF comparing estimates derived from (1) Wave 1 versus Wave 3 data, (2) 12- to 14-year-

²To examine whether we would have obtained substantively different results by using another method, we estimated with the Wave 1 data a 2-PL model using MULTILOG (Thissen et al., 2003). The item severity thresholds obtained from the 2-PL model correlated extremely highly with the severity estimates obtained in our Rasch model analyses, $r = 0.99$. Likewise, the discrimination parameters from the 2-PL analysis correlated highly with the item infit statistics ($r = -0.81$). Similar findings have been reported elsewhere when comparing Rasch models to multi-parameter models (Wright, 1995).

olds versus 15- to 18-year-olds in Wave 1, and (3) 18- to 20-year-olds versus 21- to 24-year-olds in Wave 3.

RESULTS

The prevalence of each of the 15 alcohol consumption and problems items from Wave 1 are presented in Table 1, ordered by their decreasing frequency of endorsement in the overall Wave 1 sample (aged 12 to 18). At Wave 1, 4,023 (63.2%) participants endorsed none of the items, and 3 (less than 0.1%) endorsed all of the items. Participants endorsing no items or endorsing all of the items are not informative in Rasch model analyses. They do not differentiate in their responses to the given items, and thus they contribute no information on the ordering of the items and the fit of items to the Rasch model. They are therefore not included in the calculation of statistics based upon Rasch model analyses. Consequently, the percentages shown in Tables 1–4 are based on the full sample, but the Rasch model estimates are based only on participants who endorsed at least 1 but fewer than all items. These differences in sample sizes are denoted in each table per each analysis, where n denotes the overall sample size for the group being analyzed and n_m denotes the number of participants whose scores contributed to the Rasch model.

Unidimensionality and Local Independence

To establish unidimensionality of the items, principal components analyses were performed on tetrachoric correlations among 12 of the 15 items. Four items (“daily,” “weekly,” “monthly,” and “12 or more drinks”) were subsets of each other and therefore could not produce valid tetrachoric correlations. Therefore, the analysis was conducted using only the “weekly” item. Analyses conducted using any of the remaining 3 consumption items in place of the “weekly” item yielded equivalent results. A strong first factor accounted for 71.1% of the variance (eigenvalue = 8.5), with the second factor only accounting for 5.6% of the variance (eigenvalue = 0.7). Likewise, parallel analysis (Horn, 1965) supported retention of only 1 factor.

Principal components analysis of residual variance after fitting the Rasch model to the 15 items supported the local independence of this set, with the first 2 residual components accounting for 1.92 and 1.32 units of variance, respectively (total item variance was 15). The first component was above the cutoff for significance of 1.50 but had high loadings (i.e., greater than 0.40) for only 2 items, “monthly” (0.79), and “weekly” (0.79). Given that these items were coded from the same question and are subsets of one another, this finding is not surprising. We retained both of these items so that we could examine where these levels of drinking frequency were placed along the alcohol involvement continuum. The other alcohol consumption items (“12 or more drinks in a year,” “5 drinks in a row,” and “daily”) did not load significantly on this factor. Analyses of residuals within sex, race/ethnicity, and age subgroups were consistent with the overall analysis.

Item-Level Estimates of Problem Severity

Table 1 presents each item's Rasch model severity estimate and standard error, infit value, and point biserial correlation between the item and the sum of the remaining items. The severity estimates in a Rasch analysis are expressed in equal interval logit units, so that the distance between –1.00 and –2.00, for example, is equivalent to the distance between 0.00 and 1.00. These estimates indicate the point along the latent continuum of alcohol involvement at which the specific item is making discriminations and indicate how similar or dissimilar items are in respect to the underlying latent continuum. Items covered a broad range of severity ranging from –2.65 (“12 or more drinks”) to 3.51 (“daily”). Items were closely spaced toward the middle of the scale relative to the high and low ends of the scale. For example, “hung over” (–0.96) and “threw up” (–0.86) had very similar severity estimates, as did “physical fight” (1.13)

and “friend problem” (1.14). The associated standard errors estimate the precision of the severity estimates. Generally, as item estimates become more extreme, their standard errors tend to be relatively higher. This is due to the fact that more severe items are observed in the profiles of a few rare cases, and thus fewer profiles are available for the estimation of these points.

For the 15 items, the infit values ranged from 0.83 to 1.10, well within the accepted range, indicating good model fit. Of the 15 items, only 1 item exhibited a low point biserial correlation (“daily,” $r = 0.15$). Given that this item was at the very extreme of the measured range of severity for this population and therefore has low variability, a low point biserial is expected as responses to this item are likely to explain relatively little variance in the rest of the items. All other items exceeded point biserial values of 0.25, indicating strong relations to the underlying continuum.

DIF

Also presented in Table 1 are indices of DIF observed in the Wave 1 sample with regard to sex. Specifically, Table 1 presents the rates of endorsement and severity estimates by sex, the DIF per item (i.e., the difference in severity estimates), and whether the item showed meaningful (i.e., >0.50 logits) and statistically significant DIF. For sex, 2 items showed significant and meaningful DIF. Getting into a physical fight while drinking (DIF = -0.51) was more severe for females compared to males, whereas having a dating problem due to drinking (0.55) was more severe for males.

Ethnic groups comparisons are presented in Table 2. In comparing non-Hispanic Whites to Blacks, significant DIF was found for 4 items. Getting drunk (-0.82) and drinking 5 drinks in a row (-0.90) were more severe for Blacks, although these items were the second and third most commonly endorsed items in both groups. Weekly drinking (0.66) and daily drinking (1.13) were more severe for Whites. The daily item had extremely low rates of endorsement, so DIF on this item is subject to substantial error. Whites differed from Hispanics only in that getting drunk was more severe for Hispanics (-0.59). Compared to Blacks, drinking 5 drinks in a row was less severe for Hispanics (0.85), whereas weekly drinking was more severe (-0.58).

Table 3 presents the prevalence of each item at Wave 3 and the Rasch model results. All items fit the Rasch model well. Table 3 also shows tests of DIF when comparing results obtained at Wave 1 (12 to 18 years old) with those obtained at Wave 3 (18 to 24 years old).³ Due to questionnaire changes made between Wave 1 and Wave 3 data collection, only 13 items were compared in this way. Eight meaningful differences were found. Generally, lower severity items, such as “12 or more drinks” (1.28), “hung over” (0.67), “monthly” drinking (1.09), and “weekly” drinking (0.80) were more severe for the respondents at a younger age (12 to 18 years) than at an older age (18 to 24 years). By contrast, alcohol problem items, such as “date problems” (-1.09), “physical fight” (-1.07), “friend problem” (-0.90), and “school or work problem” (-0.52) were estimated to be more severe at an older age.

To provide a more fine-grained analysis of age differences in item functioning, we also conducted cross-sectional comparisons, where we compared 12- to 14-year-olds with 15- to 18-year-olds in Wave 1, and 18- to 20-year-olds with 21- to 24-year-olds in Wave 3.⁴

³To determine whether results were affected by the fact that additional subjects were included in the Wave 3 analyses who were not included in the Wave 1 analysis (i.e., those who had minimal or no alcohol consumption at Wave 1), we re-ran the Wave 3 analyses with only those participants who were also in the Wave 1 Rasch model analysis ($n = 1293$). The item severity estimates obtained with this Wave 3 subsample were highly concordant with those obtained in the full Wave 3 sample with the mean absolute difference between item estimates across the analyses being 0.08 logits, range = 0.00 to 0.22.

Endorsement rates and DIF information are presented in Table 4. In comparing smaller age ranges, fewer DIFs were identified. Nonetheless, there were some. For Wave 1, 2 DIFs were identified. “Gotten drunk” (0.60) was more severe in the younger group (12 to 14 years), and “school problems” (−0.82) was more severe in the older group. For Wave 3, three DIFs were identified. Both “monthly” (0.75) and “weekly” (0.62) drinking were estimated to be more severe in the younger group (18 to 20 years), whereas “drunk at school or work” (−0.69) was more severe in the older group (21 to 24 years).

CONCLUSIONS

The present study drew on Add Health data to conduct Rasch model analyses of alcohol consumption and problems during adolescence and young adulthood. We explored the extent to which item performance differed across demographic groups and as a function of developmental stage. Below, we describe findings and offer potential explanations for subgroup differences.

The Alcohol Involvement Continuum

There was strong support for a single latent dimension underlying participant responses, consistent with other recent analyses that have attempted to combine indicators of alcohol consumption and alcohol use disorder criteria in adults (Kahler et al., 2008; Krueger et al., 2004; Saha et al., 2007). Analysis of residuals after fitting the items to the Rasch model suggested that there was little meaningful correlation among the item residuals, with the one exception of weekly and monthly drinking, the first one of which is subset of the other. That the data fit a Rasch model well provides direct evidence that measures of consumption and problems can index case severity in an additive manner across a range of alcohol involvement and across both adolescents and young adults. That does not, however, preclude future examinations of potential subgroups within the continuum using techniques such as latent class, latent trajectory, or factor mixture modeling analyses.

All 15 items examined appeared to contribute to distinguishing levels of alcohol involvement among adolescents with the one exception of daily drinking. Daily or almost daily drinking is an extremely rare occurrence among adolescents and young adults (in the present study, 0.9% and 2.0%, respectively) and does not appear to contribute meaningfully to the assessment of alcohol involvement in this general population sample, although this item may be important for clinical youth samples or for epidemiological studies exploring the benefits of frequent alcohol consumption on, for example, coronary heart disease risk among adults (Tolstrup et al., 2006). Importantly, a single alcohol consumption item, frequency of use, captured a wide range of drinking behavior. The present study suggests that no single operationalization of heavy alcohol consumption is necessary, consistent with research addressing different definitions of heavy episodic drinking (Jackson, 2008; Jackson and Sher, 2008). Different indices of consumption appear to map different regions of the underlying continuum.

Although we found evidence of unidimensionality across indices of alcohol use and problems, it is not our intent to contradict the literature suggesting that alcohol consumption and alcohol-related problems are separate dimensions that are correlated only moderately (Bailey and Rachal, 1993; DeCourville and Sadava, 1997; Sadava and Galanter, 1990; Stice et al., 1998) and may have different correlates (Stice et al., 2002; Windle, 1996). Rather, we are suggesting that because alcohol consumption and alcohol problems tap different constructs within a given

⁴We also compared the smaller age range groups longitudinally. Namely, we compared the item severity estimates for participants aged 12 to 14 years of age during Wave 1 with their estimates when they were aged 18 to 20 years of age during Wave 3, and participants aged 15 to 18 years of age during Wave 1 with their estimates when they were aged 21 to 24 years of age during Wave 3. The results were consistent with those presented in Table 4, and thus are not presented in further detail.

domain, integrating them in a single dimension may be a useful way to construct a continuum of adolescent alcohol involvement that is sensitive to distinctions in drinking behavior (Windle, 1996).

Subgroup Differences

A few of the items showed DIF by sex and race/ethnicity. Consistent with the broader literature (Wagner et al., 2002; Wilsnack et al., 2000), boys had slightly heavier consumption rates and greater endorsement of some problems than girls. Yet, girls were equally likely to endorse more social consequences of alcohol use, consistent with Perkins (1992). However, despite the many sex differences in item prevalence, only 2 items performed substantially differently by sex: physical fights and dating problems. Physical fights was a more severe item for girls than for boys, consistent with prior research with college students (Kahler et al., 2004), and likely reflects the higher frequency of physical fighting in males regardless of alcohol consumption. Dating problems were a less severe item for girls than boys. That dating problems due to drinking were more common for girls than for boys given the same level of alcohol involvement indicates that girls may be especially prone to alcohol's affecting romantic relationships. Given the link between early pubertal timing and early substance use (Dick et al., 2000; Tschann et al., 1994), it is possible that romantic relationships are occurring in younger girls, who are less able to consume alcohol in a safe manner.

Although not hypothesized, getting drunk and drinking 5 drinks in a row showed different performance across Whites, Blacks, and Hispanics. These were less severe items for White adolescents with drinking 5 drinks in a row being particularly severe for Black adolescents consistent with the broader literature showing heavier drinking among Whites (Johnston et al., 2006). At the same time, for all groups, getting drunk and drinking 5 drinks in a row were the second and third most commonly endorsed items and fit a Rasch measurement model well. These items are likely to be endorsed by all adolescents who report any mild (e.g., having a hangover) or more severe (e.g., being in a regretted sexual situation) alcohol-related problems. That daily drinking showed significant DIF by race/ethnicity further indicates that this item may have limited utility for measurement. Weekly drinking showed DIF as well, being less severe for Black adolescents. Given these findings, dichotomizing drinking frequency using a cut-off of monthly drinking appears most useful for measurement. Using only monthly drinking, but not weekly drinking, would ensure that the index does not contain locally dependent items.

Greater differences in item performance were observed across age, which is perhaps not surprising given that the prevalence of most alcohol involvement items increases throughout adolescence and young adulthood as consumption becomes more normative. However, the differences in item functioning indicated that the prevalence of certain items increased more so than other items and the relative severity of most of the items changed. In particular, a number of indices of consumption and acute alcohol effects (12 or more drinks in a year, being hung over, drinking monthly, and drinking weekly) became less severe indicators of alcohol involvement in young adulthood. These items were all indicators of relatively low severity. By contrast, the more severe indicators of alcohol involvement at Wave 1 became even more severe at Wave 3. These findings reflect an important pattern in the data. The differences in item locations between consumption items and alcohol problems were generally smaller when participants were adolescents than when they were young adults. This indicates, for example, that adolescents who drink weekly are significantly more likely to experience a problem with a friend or have a physical fight while drinking than are young adults who drink weekly. In fact, the rate of weekly drinking in adolescents was only marginally higher than the rates of a number of psychosocial alcohol problems. Even within the Wave 3 sample, weekly and monthly drinking were significantly less severe indicators of involvement among those of legal

drinking age compared to those below the legal drinking age. Although drinking frequency measures may be relatively good indicators of alcohol involvement severity in adolescents, the meaning of these indicators necessarily changes with increasing age. These results highlight the fact that although measures of consumption and problems can be modeled on a single latent continuum, they also are distinct constructs that may show differential associations with the continuum during development.

Although drinking frequency indices showed differential functioning by age, drinking 5 drinks in a row performed similarly across all age comparisons. These results are consistent with previous studies that have found that the 5+ drinks measure discriminates well between those who do and those who do not show symptoms of alcohol abuse and dependence (Saha et al., 2007) and is predictive of subsequent alcohol problems (Jackson, 2008; Midanik et al., 1996; Wechsler et al., 1994) as well as alcohol dependence (Knight et al., 2002). This item may serve as a good “anchor” item when examining the risk of expressing certain alcohol problems and symptoms of alcohol use disorders among high risk individuals across the lifespan, but the differential severity of this item for Black adolescents should be considered as noted above. The relative distances, expressed in logit units, between heavy drinking and specific alcohol involvement symptoms indicates how likely individuals in a given subgroup are to experience alcohol problems given that they have drunk alcohol at a level likely to produce intoxication.

The drift in the severity of specific indicators of alcohol involvement with increasing age may reflect a number of processes. First, individuals who are drinking at a young age may be especially prone to experiencing alcohol-related problems because they are marked by particular characteristics that put them at risk, such as having conduct disorder or delinquent peers. (However, at Wave 3, the relative ordering of items did not differ when we restricted analyses to those participants who also had significant alcohol involvement at Wave 1.) It also may be the case that alcohol affect the brains of adolescents differently (Brown et al., 2000; Monti et al., 2005), and adolescents may have less cognitive control capacity for avoiding alcohol-related negative consequences when they drink (Brown et al., 2008; Kirisci et al., 2004). Environmental contexts of drinking also change with age. In young adolescents, drinking alcohol may be a more deviant behavior and therefore less likely to be sanctioned by peers and parents, whereas alcohol involvement is more normative in young adulthood (Maggs and Schulenberg, 2004-2005). Based on the difference observed between legal-age and below legal-age young adults in the alcohol involvement severity indicated by monthly and weekly drinking, it does appear that regular drinking that is illegal is especially likely to be associated with consequences.

Other developmental factors also may be relevant, as has been highlighted by Neal and colleagues (2006). For example, among young adults, being drunk at school or work was less severe for those 18 to 20 years of age than for those 21 to 24 years of age. This may reflect the fact that those who are aged 21 to 24 years may be moving into the work force, and intoxication at work may be a more deviant behavior than being drunk while at class in college. This is consistent with work showing that heavy alcohol consumption interferes with the transition to the workforce (Schulenberg et al., 2003) but does not have a reliable negative impact on academic performance (Pascarella et al., 2007), with some studies indicating no association between heavy drinking and academic achievement once precollege characteristics were controlled (Paschall and Freisthler, 2003; Wood et al., 1997). As another example, getting into a regretted sexual situation while drinking was one of the few alcohol-related problems that did not become significantly more severe with increasing age. This likely reflects the fact that opportunities for sexual activity increase with age increasing the likelihood that unwanted sexual situations will occur when drinking occurs.

The developmental perspective that we have taken in the present study should be extended in future work. For example, adolescents who have a first drinking experience at a young age, perhaps due to greater social or physical availability, may show different symptom profiles when compared to later-onsetting youth. The next step would be to examine the extent to which we observe DIF as a function of different drinking courses, defined either a priori using age of onset or based on growth in heavy drinking, or by using mixture modeling techniques (e.g., comparing those in an early onsetting heavy course vs. those in a later-onsetting moderate course). It is possible that not only do items perform differently across developmental stage, but they also differ across the course of alcohol involvement taken after first exposure to alcohol (i.e., time defined not by age but by time since first drink).

STRENGTHS AND LIMITATIONS

The present study drew on a large general population sample of adolescents and young adults, and in contrast to much research on emerging adulthood, was not limited to college student samples. Given that college students are more likely to drink heavily but to consume alcohol less frequently than their non-college-bound peers (Slutske, 2005), studies examining the alcohol involvement continuum in emerging adulthood should not limit respondents to college attenders. The prospective design permitted examination of the transition from adolescence to young adulthood, and the multi-cohort design allowed examination of cross-sectional age differences.

There were several limitations that should be noted. There was considerable attrition between Waves 1 and 3, which was related to demographic factors as well as alcohol involvement. Analyses were limited to the relatively small set of alcohol consumption and problem-related items included in the Add Health dataset. There were no items that were designed specifically to assess alcohol use disorder criteria; therefore, it is not clear from the present study that how alcohol use disorder criteria would function in relation to the items that were analyzed. In addition, the same item set was not retained across time, with 2 items dropped and one item added at Wave 3.

CONCLUSIONS

In summary, this study found that indices of alcohol consumption and alcohol-related problems or negative consequences can be combined in a single index when examining adolescent drinking. This index is relatively invariant in its performance across sexes and ethnicities and may therefore serve as a valuable dependent variable when examining risk factors for change in alcohol involvement during adolescence. In particular, the rank order of the prevalence of the symptoms remains quite stable during adolescence. However, the transition from adolescence to young adulthood altered the performance of a number of the items in a manner that is consistent with expected developmental changes. Scores on a summative index obtained in adolescence can be compared to those obtained in young adulthood, as both are valid measures. However, the interpretation of the scores must be attentive to developmental and environmental differences. The number of individuals scoring in the lower regions of the continuum are likely to increase markedly as drinking increases, while the number of individuals scoring in the high ranges indicative of significant alcohol problems will increase more slowly. Greater understanding of the factors that contribute to these developmental differences may ultimately inform efforts to prevent negative consequences of drinking among youth.

Acknowledgments

This study was supported by grant R21 AA016524 from the National Institute on Alcohol Abuse and Alcoholism.

REFERENCES

- Baer JS, Kivlahan DR, Marlatt GA. High-risk drinking across the transition from high school to college. *Alcohol Clin Exp Res* 1995;19:54–61. [PubMed: 7771663]
- Bailey SL, Rachal JV. Dimensions of adolescent problem drinking. *J Stud Alcohol* 1993;54:555–565. [PubMed: 8412145]
- Bond, TG.; Fox, CM. *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. Erlbaum; Mahway, NJ: 2001.
- Brown SA, McGue M, Maggs J, Schulenberg J, Hingson R, Swartzwelder S, Martin C, Chung T, Tapert SF, Sher K, Winters KC, Lowman C, Murphy S. A developmental perspective on alcohol and youths 16 to 20 years of age. *Pediatrics* 2008;121(Suppl. 4):S290–S310. [PubMed: 18381495]
- Brown SA, Tapert SF, Granholm E, Delis DC. Neurocognitive functioning of adolescents: effects of protracted alcohol use. *Alcohol Clin Exp Res* 2000;24:164–171. [PubMed: 10698367]
- Chung T, Martin CS, Winters KC. Diagnosis, course, and assessment of alcohol abuse and dependence in adolescents. *Recent Dev Alcohol* 2005;17:5–27. [PubMed: 15789857]
- DeCourville NH, Sadava SW. The structure of problem drinking in adulthood: Aconfirmatory approach. *J Stud Alcohol* 1997;58:146–154. [PubMed: 9065892]
- Dick DM, Rose RJ, Viken RJ, Kaprio J. Pubertal timing and substance use: associations between and within families across late adolescence. *Dev Psychol* 2000;36:180–189. [PubMed: 10749075]
- Draba, RE. The identification and interpretation of item bias.. Educational Statistics Laboratory, memo 25, University of Chicago. 1977. Available at: <http://www.rasch.org/memo25.htm>
- Grunbaum JA, Kann L, Kinchen S, Ross J, Hawkins J, Lowry R, Harris WA, McManus T, Chyen D, Collins J. Youth risk behavior surveil-lance – United States, 2003. *MMWR Surveill Summ* 2004;53:1–96. [PubMed: 15152182]
- Hochberg Y. A sharper Bonferroni procedure for multiple tests of significance. *Biometrika* 1988;75:800–802.
- Holland, PW.; Wainer, H. *Differential Item Functioning*. Lawrence Erlbaum; Hillsdale, NJ: 1993.
- Holm S. A simple sequentially rejective multiple test procedure. *Scand J Statist* 1979;6:65–70.
- Horn JL. A rationale and test for the number of factor in factor analysis. *Psychometrika* 1965;30:179–185. [PubMed: 14306381]
- Jackson KM. Heavy episodic drinking: determining the predictive utility of five or more drinks. *Psychol Addict Behav* 2008;22:68–77. [PubMed: 18298232]
- Jackson KM, Sher KJ. Comparison of longitudinal phenotypes based on alternate heavy drinking cut scores: a systematic comparison of trajectory approaches III. *Psych Add Behav* 2008;22(2):198–209.
- Johnston, LD.; O'Malley, PM.; Bachman, JG.; Schulenberg, JE. *Monitoring the Future National Survey Results on Drug Use, 1975–2005. Volume I: Secondary school students* (NIH Publication No. 06-5883). National Institute on Drug Abuse; Bethesda, MD: 2006.
- Kahler CW, Strong DR. A Rasch model analysis of DSM-IV Alcohol abuse and dependence items in the National Epidemiological Survey on Alcohol and Related Conditions. *Alcohol Clin Exp Res* 2006;30:1165–1175. [PubMed: 16792564]
- Kahler CW, Strong DR, Hayaki J, Ramsey SE, Brown RA. An item response analysis of the Alcohol Dependence Scale in treatment-seeking alcoholics. *J Stud Alcohol* 2003a;64:127–136. [PubMed: 12608493]
- Kahler CW, Strong DR, Papandonatos GD, Colby SM, Clark MA, Boergers J, Niaura R, Abrams DB, Buka SL. Cigarette smoking and the lifetime alcohol involvement continuum. *Drug Alcohol Depend* 2008;93:111–120. [PubMed: 17964082]
- Kahler CW, Strong DR, Read JP. Toward efficient and comprehensive measurement of the alcohol problems continuum in college students: the brief young adult alcohol consequences questionnaire. *Alcohol Clin Exp Res* 2005;29:1180–1189. [PubMed: 16046873]
- Kahler CW, Strong DR, Read JP, Palfai TP, Wood MD. Mapping the continuum of alcohol problems in college students: a Rasch model analysis. *Psychol Addict Behav* 2004;18:322–333. [PubMed: 15631604]

- Kahler CW, Strong DR, Stuart GL, Moore TM, Ramsey SE. Item functioning of the alcohol dependence scale in a high-risk sample. *Drug Alcohol Depend* 2003b;72:183–192. [PubMed: 14636973]
- Kirisci L, Tarter RE, Vanyukov M, Reynolds M, Habeych M. Relation between cognitive distortions and neurobehavior disinhibition on the development of substance use during adolescence and substance use disorder by young adulthood: a prospective study. *Drug Alcohol Depend* 2004;76:125–133. [PubMed: 15488336]
- Knight JR, Wechsler H, Kuo M, Seibring M, Weitzman ER, Schuckit MA. Alcohol abuse and dependence among U.S. college students. *J Stud Alcohol* 2002;63:263–270. [PubMed: 12086126]
- Krueger RF, Nichol PE, Hicks BM, Markon KE, Patrick CJ, Lacono WG, McGue M. Using latent trait modeling to conceptualize an alcohol problems continuum. *Psychol Assess* 2004;16:107–119. [PubMed: 15222807]
- Lai JS, Teresi J, Gershon R. Procedures for the analysis of differential item functioning (DIF) for small sample sizes. *Eval Health Prof* 2005;28:283–294. [PubMed: 16123258]
- Li TK, Hewitt BG, Grant BF. The alcohol dependence syndrome, 30 years later: a commentary. The 2006 H. David Archibald lecture. *Addiction* 2007;102:1522–1530. [PubMed: 17680851]
- Linacre JM. Detecting multidimensionality: Which residual data-type works best. *J Outcome Meas* 1998;2:266–283. [PubMed: 9711024]
- Linacre JM, Wright BD. Reasonable mean-square fit values. *Rasch Meas Transac* 1994;8:370.
- Linacre, JM.; Wright, BD. A User's Guide to BIGSTEPS: A Rasch-Model Computer Program. MESA Press; Chicago, IL: 1998.
- Maggs JL, Schulenberg J. Trajectories of alcohol use during the transition to adulthood. *Alcohol Res Health* 2004-2005;28:195–201.
- Martin CS, Chung T, Kirisci L, Langenbucher JW. Item response theory analysis of diagnostic criteria for alcohol and cannabis use disorders in adolescents: implications for DSM-V. *J Abnorm Psychol* 2006;115:807–814. [PubMed: 17100538]
- Midanik LT, Tam TW, Greenfield TK, Caetano R. Risk functions for alcohol-related problems in a 1988 US national sample. *Addiction* 1996;91:1427–1437. discussion 1439–56. [PubMed: 8917925]
- Monti PM, Miranda R Jr, Nixon K, Sher KJ, Swartzwelder HS, Tapert SF, White A, Crews FT. Adolescence: booze, brains, and behavior. *Alcohol Clin Exp Res* 2005;29:207–220. [PubMed: 15714044]
- Neal DJ, Corbin WR, Fromme K. Measurement of alcohol-related consequences among high school and college students: application of item response models to the Rutgers Alcohol Problem Index. *Psychol Assess* 2006;18:402–414. [PubMed: 17154761]
- Pascarella ET, Goodman KM, Seifert TA. College student binge drinking and academic achievement: A longitudinal replication and extension. *J College Stud Dev* 2007;48:715–727.
- Paschall MJ, Freisthler B. Does heavy drinking affect academic performance in college? Findings from a prospective study of high achievers. *J Stud Alcohol* 2003;64:515–519. [PubMed: 12921193]
- Perkins HW. Gender patterns in consequences of collegiate alcohol abuse: a 10-year study of trends in an undergraduate population. *J Stud Alcohol* 1992;53:458–462. [PubMed: 1405638]
- Rasch, G. Probabilistic Models for Some Intelligence and Attainment Test. Denmarks Paedagogiske Institut; Copenhagen: 1960.
- Sadava, SW.; Galanter, M. Problem drinking and alcohol problems: widening the circle of covariation. In: Galanter, M., editor. *Recent Developments in Alcoholism, Vol. 8: Combined Alcohol and Other Drug Dependence*. Plenum Press; New York: 1990. p. 173-201.
- Saha TD, Stinson FS, Grant BF. The role of alcohol consumption in future classifications of alcohol use disorders. *Drug Alcohol Depend* 2007;89:82–92. [PubMed: 17240085]
- Schulenberg, JE.; Maggs, JL.; O'Malley, PM. How and why the understanding of developmental continuity and discontinuity is important: The sample case of long-term consequences of adolescent substance use. In: Mortimer, JT.; Shanahan, MJ., editors. *Handbook of the Life Course*. Kluwer Academic/Plenum Publishers; New York: 2003. p. 413-436.
- Sher KJ, Rutledge PC. Heavy drinking across the transition to college: predicting first-semester heavy drinking from precollege variables. *Addict Behav* 2007;32:819–835. [PubMed: 16860940]

- Slutske WS. Alcohol use disorders among US college students and their non-college-attending peers. *Arch Gen Psychiatry* 2005;62:321–327. [PubMed: 15753245]
- Smith, RM.; Miao, CY. Assessing unidimensionality for the Rasch measurement. In: Wilson, M., editor. *Objective Measurement: Theory and Practice*. Vol. 2. Ablex; Norwood, NJ: 1994. p. 316-327.
- Stice E, Barrera M, Chassin L. Prospective differential prediction of adolescent alcohol use and problem use: Examining the mechanisms of effect. *J Abnorm Psychol* 1998;107:616–628. [PubMed: 9830249]
- Stice E, Kirz J, Borbely C. Disentangling adolescent substance use and problem use within a clinical sample. *J Adolesc Res* 2002;17:122–142.
- Thissen, D.; Chen, WH.; Bock, D. *MULTILOG for Windows*, 7.03 ed. Scientific Software International, Inc; Lincolnwood, IL: 2003.
- Tolstrup J, Jensen MK, Tjonneland A, Overvad K, Mukamal KJ, Gronbaek M. Prospective study of alcohol drinking patterns and coronary heart disease in women and men. *BMJ* 2006;332:1244–1248. [PubMed: 16672312]
- Tristan A. An adjustment for sample size in DIF Analysis. *Rasch Meas Transac* 2006;20:1070–1071.
- Tschann JM, Adler NE, Irwin CE Jr, Millstein SG, Turner RA, Kegeles SM. Initiation of substance use in early adolescence: the roles of pubertal timing and emotional distress. *Health Psychol* 1994;13:326–333. [PubMed: 7957011]
- Udry, JR. The National Longitudinal Study of Adolescent Health (Add Health), Waves I & II, 1994–1996; Wave III, 2001–2002 [Machine-Readable Data File and Documentation]. Carolina Population Center, University of North Carolina at Chapel Hill; Chapel Hill, NC: 2003.
- Wagner EF, Lloyd DA, Gil AG. Racial/ethnic and gender differences in the incidence and onset age of DSM-IV alcohol use disorder symptoms among adolescents. *J Stud Alcohol* 2002;63:609–619. [PubMed: 12380858]
- Wechsler H, Davenport A, Dowdall G, Moeykens B, Castillo S. Health and behavioral consequences of binge drinking in college. A national survey of students at 140 campuses. *JAMA* 1994;272:1672–1677. [PubMed: 7966895]
- White HR, Labouvie EW. Towards the assessment of adolescent problem drinking. *J Stud Alcohol* 1989;50:30–37. [PubMed: 2927120]
- Wilsnack RW, Vogeltanz ND, Wilsnack SC, Harris TR, Ahlstrom S, Bondy S, Csemy L, Ferrence R, Ferris J, Fleming J, Graham K, Greenfield T, Guyon L, Haavio-Mannila E, Kellner F, Knibbe R, Kubicka L, Loukomskaja M, Mustonen H, Nadeau L, Narusk A, Neve R, Rahav G, Spak F, Teichman M, Trocki K, Webster I, Weiss S. Gender differences in alcohol consumption and adverse drinking consequences: cross-cultural patterns. *Addiction* 2000;95:251–265. [PubMed: 10723854]
- Windle M. An alcohol involvement typology for adolescents: Convergent validity and longitudinal stability. *J Stud Alcohol* 1996;57:627–637. [PubMed: 8913994]
- Wood PK, Sher KJ, Erickson DJ, DeBord KA. Predicting academic problems in college from freshman alcohol involvement. *J Stud Alcohol* 1997;58:200–210. [PubMed: 9065898]
- Wright BD. 3PL or Rasch? *Rasch Meas Transac* 1995;9:408–409.
- Wright BD. Local dependency, correlations and principal components. *Rasch Meas Transac* 1996;10:509–511.
- Wright, BD.; Masters, GN. *Rating Scale Analysis*. MESA; Chicago: 1982.

Table 1
Items, Rates of Past 12-Month Endorsement, Item Response Model Parameters, and Differential Item Functioning by Sex

	Total sample statistics					Comparison by sex					
	Item content (least to most severe)	% (<i>n</i> = 6,353)	EST (<i>n_m</i> = 2,327)	Model fit		Male		Female		DIF	
				SE	Infit	PTBS	% (<i>n</i> = 3,059)	EST (<i>n_m</i> = 1,157)	% (<i>n</i> = 3,294)		EST (<i>n_m</i> = 1,167)
12 or more drinks		28.3	-2.65	0.05	0.94	0.38	30.7	-2.84	26.2	-2.52	-0.32
Gotten drunk		27.0	-2.41	0.05	0.83	0.46	28.5	-2.42	25.6	-2.43	0.01
5 drinks in a row		25.6	-2.15	0.05	0.96	0.41	28.4	-2.40	23.0	-1.97	-0.43
Hung over		18.1	-0.96	0.05	1.06	0.38	18.1	-0.73	18.1	-1.19	0.46
Threw up		17.5	-0.86	0.05	1.06	0.38	17.4	-0.63	17.5	-1.11	0.48
Monthly		17.2	-0.83	0.05	0.95	0.46	19.9	-1.00	14.7	-0.68	-0.32
Regret action		13.6	-0.29	0.05	0.95	0.45	13.5	-0.05	13.8	-0.54	0.49
Weekly		9.3	0.44	0.05	0.98	0.41	11.9	0.21	6.9	0.70	-0.49
Parent problem		9.1	0.49	0.05	1.10	0.30	10.2	0.49	8.0	0.48	0.01
Date problem		8.4	0.62	0.05	1.05	0.34	8.1	0.89	8.7	0.34	0.55*
Regret sexual situation		7.9	0.73	0.05	0.98	0.39	8.2	0.87	7.5	0.57	0.30
Physical fight		6.1	1.13	0.06	0.99	0.34	8.0	0.90	4.2	1.41	-0.51*
Friend problem		6.0	1.14	0.06	1.00	0.33	5.9	1.39	6.1	0.89	0.50
School problem		3.0	2.08	0.08	0.99	0.26	3.8	2.01	2.3	2.17	-0.16
Daily		0.9	3.51	0.14	1.09	0.15	1.3	3.31	0.5	3.87	-0.56

EST, Rasch model severity estimate; DIF, differential item functioning; PTBS, point biserial correlations; *n_m*, number of participants with at least one but not all symptoms, who thereby are included in the Rasch model.

Infit values of 1 indicate perfect fit to the Rasch model; lower values indicate less random variation than expected and higher values indicate more random variation. Positive gender DIF indicates an item is more severe for males than for females.

* DIF that was statistically significant at the 0.05 level and above a moderate effect size cut-off point of 0.50.

Table 2
Items, Rates of Past 12-Month Endorsement, Item Response Model Severity Estimates, and Differential Item Functioning by Race

Item content (least to most severe)	Race statistics					Pairwise comparisons			
	White		Black		Hispanic/Latino EST (n_m = 281)	White versus Black		White versus Hispanic	
	% (n = 3,791)	EST (n_m = 1,580)	% (n = 1,507)	EST (n_m = 355)		DIF		DIF	DIF
12 or more drinks	32.9	-2.77	17.3	-2.41	29.4	-2.44	-0.36	-0.33	0.03
Gotten drunk	32.3	-2.67	14.9	-1.85	27.2	-2.08	-0.82*	-0.59*	0.23
5 drinks in a row	30.4	-2.33	13.1	-1.43	28.3	-2.28	-0.90*	-0.05	0.85*
Hung over	21.0	-0.98	10.0	-0.71	20.8	-1.12	-0.27	0.14	0.41
Threw up	20.9	-0.96	9.5	-0.60	17.2	-0.62	-0.36	-0.34	0.02
Monthly	19.7	-0.8	11.5	-1.05	16.8	-0.56	0.25	-0.24	-0.49
Regret action	16.0	-0.31	8.4	-0.34	14.1	-0.18	0.03	-0.13	-0.16
Weekly	10.1	0.57	7.4	-0.09	10.0	0.49	0.66*	0.08	-0.58*
Parent problem	10.6	0.49	4.9	0.57	11.0	0.32	-0.08	0.17	0.25
Date problem	9.3	0.7	5.3	0.47	10.8	0.34	0.23	0.36	0.13
Regret sexual situation	9.4	0.69	4.5	0.72	8.0	0.84	-0.03	-0.15	-0.12
Physical fight	6.8	1.18	3.6	1.02	7.8	0.9	0.16	0.28	0.12
Friend problem	6.8	1.18	2.9	1.33	7.9	0.87	-0.15	0.31	0.46
School problem	3.2	2.19	2.2	1.69	3.2	2.14	0.50	0.05	-0.45
Daily	0.8	3.82	0.9	2.69	1.1	3.38	1.13*	0.44	-0.69

EST, Rasch model severity estimate; DIF, differential item functioning; n_m , number of participants with at least one but not all symptoms, who thereby are included in the Rasch model.

* DIF that was statistically significant at the 0.05 level and above a moderate effect size cut-off point of 0.50.

Table 3

Wave 3 Items, Rates of Past 12-Month Endorsement, Item Response Model Parameters, and Differential Item Functioning by Assessment Year

Item content (least to most severe)	Wave 3 sample statistics				Longitudinal comparison ^d			
	% (n = 4,761)	EST (n _m = 3,094)	Model Fit		Wave 3 (18–24 years) (n _m = 3,093) EST	Wave 1 (12–18 years) (n _m = 2,298) EST	DIF	
			SE	Infit				PTBS
12 or more drinks	58.9	−4.18	0.07	1.05	0.26	−3.97	−2.69	−1.28*
Gotten drunk	50.6	−2.83	0.05	0.88	0.45	−2.63	−2.44	−0.19
5 drinks in a row	48.6	−2.57	0.05	0.98	0.43	−2.37	−2.17	−0.20
Hung over	41.6	−1.82	0.04	0.94	0.48	−1.62	−0.95	−0.67*
Threw up	34.1	−1.07	0.04	1.12	0.41	−0.88	−0.85	−0.03
Monthly	44.6	−2.10	0.05	0.97	0.47	−1.90	−0.81	−1.09*
Weekly	28.6	−0.52	0.04	0.92	0.50	−0.33	0.47	−0.80*
Date problem	10.1	1.54	0.06	0.97	0.35	1.74	0.65	1.09*
Regretted sexual situation	14.5	0.94	0.05	1.03	0.38	1.13	0.76	0.37
Physical fight	7.3	2.03	0.06	1.04	0.31	2.23	1.16	1.07*
Friend problem	8.1	1.88	0.06	0.88	0.38	2.08	1.18	0.90*
School problem	5.5	2.44	0.07	0.91	0.34	2.64	2.12	0.52*
Drunk at school or work	4.9	2.59	0.07	1.01	0.28	—	—	—
Daily	2.0	3.67	0.11	1.12	0.15	3.88	3.57	0.31

EST, Rasch model severity estimate; DIF, differential item functioning; TBS, point biserial correlations; n_{IT} , the number of participants with at least one but not all symptoms, who thereby are included in the Rasch model.

Infit values of 1 indicate perfect fit to the Rasch model; lower values indicate less random variation than expected and higher values indicate more random variation.

^a Longitudinal comparison estimates vary slightly from the estimates from the cross-sectional analyses because analyses were conducted using only those items that were common to both waves. Prevalence for each item remains unchanged and is not presented.

* DIF that was statistically significant at the 0.05 level and above a moderate effect size cut-off point of 0.50.

Items, Rates of Past 12-Month Endorsement, Item Response Model Severity Estimates, and Differential Item Functioning by Age

Item content (least to most severe)	Wave 1				Wave 3					
	12–14 years		15–18 years		DIF	18–20 years		21–24 years		
	% (n = 2,030)	EST (n _m = 402)	% (n = 4323)	EST (n _m = 1,925)		% (n = 1,627)	EST (n _m = 1,012)	% (n = 3,134)	EST (n _m = 2,082)	
12 or more drinks	14.1	–2.46	35.0	–2.69	0.23	55.0	–3.90	60.8	–4.35	0.45
Gotten drunk	12.2	–1.93	34.0	–2.53	0.60*	50.7	–3.11	50.5	–2.72	–0.39
5 drinks in a row	12.3	–1.95	31.8	–2.19	0.24	46.8	–2.58	49.5	–2.58	0.00
Hung over	9.0	–1.02	22.4	–0.94	–0.08	39.2	–1.70	42.8	–1.89	0.19
Threw up	7.1	–0.51	22.3	–0.94	0.43	34.3	–1.20	34.0	–1.01	–0.19
Monthly	7.9	–0.72	21.6	–0.85	0.13	38.5	–1.61	47.8	–2.36	0.75*
Regretted action	4.9	0.19	11.0	0.55	–0.36	–	–	–	–	–
Weekly	3.7	0.62	12.0	0.41	0.21	23.4	–0.11	31.3	–0.73	0.62*
Parent problem	6.0	–0.18	17.2	–0.31	0.13	–	–	–	–	–
Date problem	3.4	0.76	10.8	0.59	0.17	10.6	1.38	9.8	1.63	–0.25
Regret sexual situation	2.7	1.1	10.3	0.66	0.44	15.5	0.74	13.9	1.05	–0.31
Physical fight	3.1	0.89	7.4	1.17	–0.28	7.2	1.96	7.4	2.07	–0.11
Friend problem	3.0	0.96	7.4	1.18	–0.22	8.2	1.76	8.0	1.95	–0.19
School problem	2.1	1.41	3.4	2.23	–0.82*	5.2	2.40	5.6	2.46	–0.06
Drunk at school or work	–	–	–	–	–	6.2	2.17	4.1	2.86	–0.69*
Daily	0.6	2.84	1.0	3.66	–0.82	1.7	3.79	2.2	3.63	0.16

EST, Rasch model severity estimate; DIF, differential item functioning; n_{ij} , number of participants with at least one but not all symptoms, who thereby are included in the Rasch model.

* DIF that was statistically significant at the 0.05 level and above a moderate effect size cut-off point of 0.50.