Modeling the Severity of Drinking Consequences in First-Year College Women: An Item Response Theory Analysis of the Rutgers Alcohol Problem Index*

AMY M. COHN, PH.D.,† BRETT T. HAGMAN, PH.D.,† FIONA S. GRAFF, PSY.M.,† AND NORA E. NOEL, PH.D.†

Department of Mental Health Law & Policy, University of South Florida, 1330 Bruce B. Downs Boulevard, MHC 2716, Tampa, Florida 33612

ABSTRACT. Objective: The present study examined the latent continuum of alcohol-related negative consequences among first-year college women using methods from item response theory and classical test theory. **Method:** Participants (N = 315) were college women in their freshman year who reported consuming any alcohol in the past 90 days and who completed assessments of alcohol consumption and alcohol-related negative consequences using the Rutgers Alcohol Problem Index. **Results:** Item response theory analyses showed poor model fit for five items identified in the Rutgers Alcohol Problem Index. Two-parameter item response theory logistic models were applied to the remaining 18 items to examine estimates of item difficulty (i.e., severity) and discrimination parameters. The item difficulty parameters ranged from 0.591 to 2.031, and the discrimination parameters ranged from 0.321 to 2.371. Classical test theory analyses indicated that the omission of the five misfit items did not significantly alter the psychometric properties of the construct. **Conclusions:** Findings suggest that those consequences that had greater severity and discrimination parameters may be used as screening items to identify female problem drinkers at risk for an alcohol use disorder. (*J. Stud. Alcohol Drugs, 72,* 981–990, 2011)

HEAVY EPISODIC DRINKING AMONG COLLEGE women is on the rise (Wechsler et al., 2002), and the transition from high school to the first year of college may be particularly detrimental for women (Fisher et al., 2000; LaBrie et al., 2007). Women report significantly heavier rates of drinking in their freshman year relative to any other year of college (McCabe, 2002). Risk of physical and sexual violence victimization is also associated with increased alcohol consumption and heavy episodic drinking during this period (Mohler-Kuo et al., 2004; Parks et al., 2008; Testa et al., 2010). It is important to identify which consequences may be most harmful to women when heavy drinking begins (i.e., during the first year of college) so that public health campaigns and university wellness programs can intervene before the onset of an alcohol use disorder (Hingson and Kenkel, 2004; Hingson et al., 2002, 2009; U.S. Department of Health and Human Services, 2007).

Demonstrating what is termed "the telescoping effect," heavy drinking women show an accelerated rate of alcohol-related morbidity and mortality at lower levels of consumption and problem severity when compared with men (Johnson et al., 2005; Piazza, 1989). Thus, despite drinking less and for a shorter period over the life span, as a result of drinking women show greater and more rapid impairment across a number of domains that include physical illness, reproductive health, cognitive and motor performance, violence, and risky sexual behavior (Haas and Peters, 2000; Mann et al., 2005; Nolen-Hoeksema, 2004; Piazza et al., 1989). These consequences are particularly concerning given recent evidence that college women report drinking in greater amounts as a way to "keep up with" male peers and to make themselves more sexually attractive (Carpenter et al., 2008).

Problems stemming from women's heavy drinking are unique from and fall into distinct domains compared with their male counterparts (Johnston et al., 2004a, 2004b). Whereas the consequences reported by men can be classified as being physically high risk and overt (e.g., physical aggression; Neal and Carey, 2007; Neal and Fromme, 2007; Perkins, 2002; Sugarman et al., 2009), women's consequences are often less visible and typically involve damage to the self (e.g., blacking out) and sexual or physical victimization (Fisher et al., 2000). These problems are less often reported

Received: March 24, 2011. Revision: June 13, 2011.

*This research was supported in part by National Institute on Alcohol Abuse and Alcoholism Grant T32 AA07569 and National Institute on Drug Abuse (NIDA) Grant P30DA028807. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIDA or the National Institutes of Health. Portions of this research were presented at the 2010 Association for Behavioral and Cognitive Therapies convention, San Francisco, CA.

†Correspondence may be sent to Amy M. Cohn at the above address or via email at: amycohn@usf.edu. Amy M. Cohn and Brett T. Hagman are also with the Department of Criminology, University of South Florida, Tampa, FL. Fiona S. Graff is with the Center of Alcohol Studies and the Graduate School of Applied and Professional Psychology, Rutgers, The State University of New Jersey, Piscataway, NJ. Nora E. Noel is with the Department of Psychology, University of North Carolina–Wilmington, Wilmington, NC.

to friends or authorities or readily observed by others (Ullman, 2007). Heavy drinking, as opposed to social drinking, is often discouraged among women because it does not fit with traditional female gender roles (Horwitz and White, 1987; MacNeela and Bredin, 2011; Nolen-Hoeksema, 2004), and consequently women experience social impairment more readily as a result of problematic alcohol use (Dawson et al., 1995; Nolen-Hoeksema, 2004). Interpersonal problems as a result of drinking may have particularly deleterious effects on women in their first year of college because this is when they try to form new friendships and "fit in" with a peer group (LaBrie et al., 2007; MacNeela and Bredin, 2011). The next step in this line of research is to examine the severity and degree to which specific drinking consequences map onto the latent continuum of alcohol problems relevant to college women during their formative years of drinking. To do so, item-level indicators such as those derived through item response theory (IRT) statistics are needed.

The majority of studies evaluating associations of alcohol problems in college samples have not used IRT but rather have focused exclusively on classical test theory (CTT) statistics (e.g., test–retest reliability, construct validity), which present several disadvantages. CTT statistics are sample dependent and do not provide information on how well an item in a questionnaire maps onto the underlying continuum (Embretson and Reise, 2000). It is not possible to use CTT statistics to obtain information on the quality of specific items within a measure. A primary advantage of IRT is the ability to statistically model observable manifestations of a hypothesized underlying trait; in this case, alcohol problem severity (Embretson and Reise, 2000). Further, IRT can provide information about how well an item places an individual on a specific point along the alcohol problem severity continuum (Embretson and Reise, 2000; Neal et al., 2006), and because it is not sample dependent, IRT assumptions maintain that replication of findings will be consistent across similar samples (Embretson and Reise, 2000). IRT methods can also tell us which consequences reliably assess the range and severity of problems experienced by first-year college women and can assist in ascertaining which consequences may be "biased," in the sense of accurately discriminating among individuals with higher and lower levels of problem severity. IRT can also identify items that are endorsed less frequently or not at all.

Because of its strong psychometric properties and extensive application, the Rutgers Alcohol Problem Index (RAPI; White and Labouvie, 1989) has been a useful tool by which one can evaluate alcohol use consequences in a number of drinking populations, particularly college students (see Neal et al., 2006, for more detail). However, only two studies have used IRT to examine alcohol use problems in male and female college students using the RAPI. Neal et al. (2006) found that only Item 21 ("Had a bad time") was a better discriminator of high and low levels of alcohol problems for women than for men during the transition from high school to college. Earleywine et al. (2008) found that women provided higher answers on Item 4 ("Went to work or school high or drunk") at levels of alcohol problems comparable to those of men, whereas women endorsed Items 16 ("Passed out or fainted suddenly") and 18 ("Had a fight, argument, or bad feelings with a family member") at lower levels of alcohol problems.

There are several notable limitations of these IRT studies. First, the studies showed inconsistencies in their findings, with separate RAPI items that were biased (showed differential probability of endorsement) toward women (Items 4, 16, 18, 21) and toward men (Items 2, 17, 19). Further replication of item-level analysis of the RAPI among specific subpopulations is therefore warranted. In addition, Neal et al. (2006) did not examine alcohol problems that transpired over the course of the first year of college, leaving open the possibility that replicability of IRT findings may differ by examining drinking consequences after the high school transition period, once students become socialized into drinking "settings." Given that the first year of college is particularly relevant to women's heavy drinking patterns, we felt it was important to re-examine IRT analyses during this period to capture the severity of problems within this specific group.

To our knowledge, no study has applied IRT methods to examine alcohol-related consequences reported by women during their first year of college. If we can more accurately identify consequences of drinking that are of particular concern in this group, we can begin to develop targeted assessment and intervention protocols through early detection. The current study aimed to apply IRT analyses to the RAPI to capture information about the severity of drinking consequences and the distinction between those with higher and lower levels of problem severity in a sample of first-year college women. An exploratory aim was to examine the psychometric properties of a modified "re-fitted" RAPI, which was obtained after omitting poorly fitting items.

Method

Participants

Participants were 315 female undergraduate students enrolled in introductory psychology courses at a public Southeastern university. This is a secondary analysis of a study that assessed the impact of peer networks on college student drinking (Hagman et al., 2010). Data were collected during the Fall 2008, Spring 2009, and Fall 2009 semesters. Individuals were included in the current analyses if they were female, reported any alcohol use in the prior 90 days, and were in their first year of college. The mean age of the sample was 18 years (SD = 4.58); 85.1% were White, 6.8% were Hispanic or African American, and 8.1% represented other racial/ethnic groups. The majority (85.7%) lived on

campus, and most subjects (85.4%) were not involved in a sorority.

Procedure

Participants were recruited via advertisements that invited individuals to enroll in a study on attitudes toward alcohol and other drug use. After providing informed consent, respondents completed an anonymous battery of questionnaires. At the completion of the study, students were given credit as part of course requirements. All procedures were approved by the university's institutional review board.

Measures

Alcohol use. Reports of alcohol use in the past 90 days were collected via a modified quantity/frequency index. Respondents were asked to estimate their frequency and quantity of consuming beer, wine, and distilled spirits. In addition, participants were asked to report the "typical" number of drinks consumed per weekday and weekend, as well as the largest amount of alcohol consumed in a 24-hour period. An index of alcohol use frequency was created by summing the frequency of beer, wine, and distilled spirits items and an index of alcohol use quantity was created by summing the amounts of beer, wine, and distilled spirits consumed per drinking occasion.

Alcohol-related negative consequences. The RAPI (White and Labouvie, 1989) is a 23-item self-report instrument that measures frequency of alcohol-related consequences (0 = *never* to 5 = *more than 10 times*) that occurred in the past year. For the purposes of IRT analyses, response options were dichotomized (0 = *did not occur* to 1 = *did occur*) because of low endorsement for consequences that occurred more than once (e.g., 3 = *daily*). Previous studies support dichotomizing outcomes as a method to control for estimation problems that may arise when creating IRT parameter estimates with a variable that has low or infrequent endorsement (Neal et al., 2006). A dichotomously scored RAPI has been shown to be reliable and valid in college students (Martens et al., 2007).

Data analytic plan. Two-parameter IRT logistic models were applied to each of the RAPI items, which provide estimates of item difficulty parameters (values range from -3 to 3) and the slope or item discrimination parameters (values range from 0 to 3). Larger difficulty parameters provide evidence that higher values of the underlying latent trait are necessary to endorse each item; the majority of the sample is less likely to endorse an item with a large difficulty value. Larger discrimination parameters indicate that the likelihood of endorsing an item increases more rapidly as the underlying latent trait increases. An item with a large discrimination parameter is more accurate at distinguishing between those with high or low levels of alcohol problem severity.

IRT analyses were conducted via a series of steps. First, we conducted a confirmatory factor analysis on the original 23-item RAPI to ensure that assumptions of IRT were met. The first assumption, "local independence," assumes that items on the scale are independent after controlling for respondent trait level. The second assumption, "unidimensionality" (Embretson and Reise, 2000), means that items within a measure should approximate a single factor. If the assumption of unidimensionality is met, then the assumption of local independence is met (Embretson and Reise, 2000). Because the RAPI items were dichotomized, we specified a single-factor model with tetrachoric correlations.

Next, an item "misfit" analysis was conducted to examine the extent to which each item of the original RAPI fit the specified two-parameter model (i.e., difficulty and discrimination). When a specific item does not fit a chosen model, estimated parameters may be less accurate and should be considered for removal. Thus, item misfit analyses examine the overall performance of each item across specific subgroups of "performance" (e.g., higher/low alcohol use severity), and those that show poor fit can be dropped from the model (Stone and Zhang, 2003). For this step, respondents are grouped into intervals based on their overall ability and compared with an expected distribution of ability levels (Stone and Zhang, 2003). Chi-square tests of significance were used to determine item misfit, with statistically significant values indicating misfit. To control for Type 1 error inflation, the overall α was set at $p < .001$.

Item characteristic curves were plotted and examined for each item that demonstrated misfit. Item characteristic curves model the probability that an item is endorsed as a function of the value of the purported underlying latent trait (e.g., severity of alcohol use consequences). The typical item characteristic curve has a well-defined S-shape and indicates that the probability of endorsing a specific item increases monotonically as the latent trait increases.

A final IRT analysis (i.e., difficulty and discrimination parameters) was used to examine those remaining items that did not demonstrate significant item misfit. CTT psychometric analyses were computed on this final set using the original RAPI scoring scheme (0–5). Internal consistency was assessed by calculating Cronbach's α with each item removed, as well as the item to total scale correlations. Associations with alcohol consumption variables were also conducted to examine the construct validity of the "re-fitted" RAPI. An exploratory principal components analysis was then conducted on the original and modified RAPI, with the goal of validating whether a modified RAPI, after omitting poorly fitting items, remained similar in construct validity to the original RAPI and accounted for a roughly equal amount of variance. This approach has been used in prior IRT studies of alcohol and drug use consequences (Hagman et al., 2009).

Last, total information curves for the original RAPI and a modified RAPI were plotted. These were estimated by all

values from the difficulty and discrimination parameters and indicate the amount of information the scale provides across the underlying latent-trait continuum (i.e., the point on the latent-trait continuum at which the scale is most reliable). IRT models were analyzed using PARSCALE 4.1 (Muraki and Bock, 2002) with a Bayesian expectation-maximization estimation procedure. The criterion for convergence of the expectation-maximization estimation procedure was 0.005.

Results

Drinking characteristics

Participants drank, on average, on 18% of the days (*SD* = 17.56). Weekly consumption of beer (26.3%; *n* = 116) was greater than that of either distilled spirits (19.8%; *n* = 78) or wine (7.9%; *n* = 25). The average number of standard drinks consumed per drinking day was 2.17 (*SD* = 2.65) for weekdays and 4.94 (*SD* = 4.71) for weekend days.

Item response theory assumptions

Because prior research suggests that the RAPI may have more than one factor when scored dichotomously (Martens et al., 2007), we first conducted an exploratory factor analysis to examine the factor structure of the RAPI in this population. A two-factor solution emerged, accounting for 63.11% (eigenvalue = 14.516) and 7.35% (eigenvalue = 1.52) of the variance for the first and second factors, respectively, which had a correlation of $r = .54$. Model fit indices for a two-factor confirmatory factor analysis indicated slightly better model fit than a one-factor solution: comparative fit index (CFI) = .977; Tucker–Lewis index (TLI) = .974; and root mean square error of approximation (RMSEA) = .046.

We then implemented an approach similar to that of Neal et al. (2006), in which a single-factor solution was specified for the RAPI using confirmatory factor analysis to assess the assumption of unidimensionality. Results of the one-factor confirmatory factor analysis on the dichotomized 23-item RAPI indicated that IRT assumptions were met, with good model fit: TLI = .978, CFI = .952, RMSEA = .076.

Item misfit analysis

The two-parameter model fit the data reasonably well, except for the following items: Item 3 ("Missed out on other things because you drank too much alcohol"), $\chi^2(1, N = 315) = 28.62$; Item 6 ("Neglected your responsibilities"), $\chi^2(1, N = 315) = 37.95$; Item 7 ("Relatives avoided you"), $\chi^2(1, N = 315) = 24.63$; Item 21 ("Had a bad time"), $\chi^2(1, N = 315) = 31.32$; and Item 23 ("Was told by a friend or neighbor to stop or cut down drinking"), $\chi^2(1, N = 315) = 29.69$ (all ps < .0001). We dropped those five items and conducted item misfit analyses again on the final set of 18 RAPI items (i.e.,

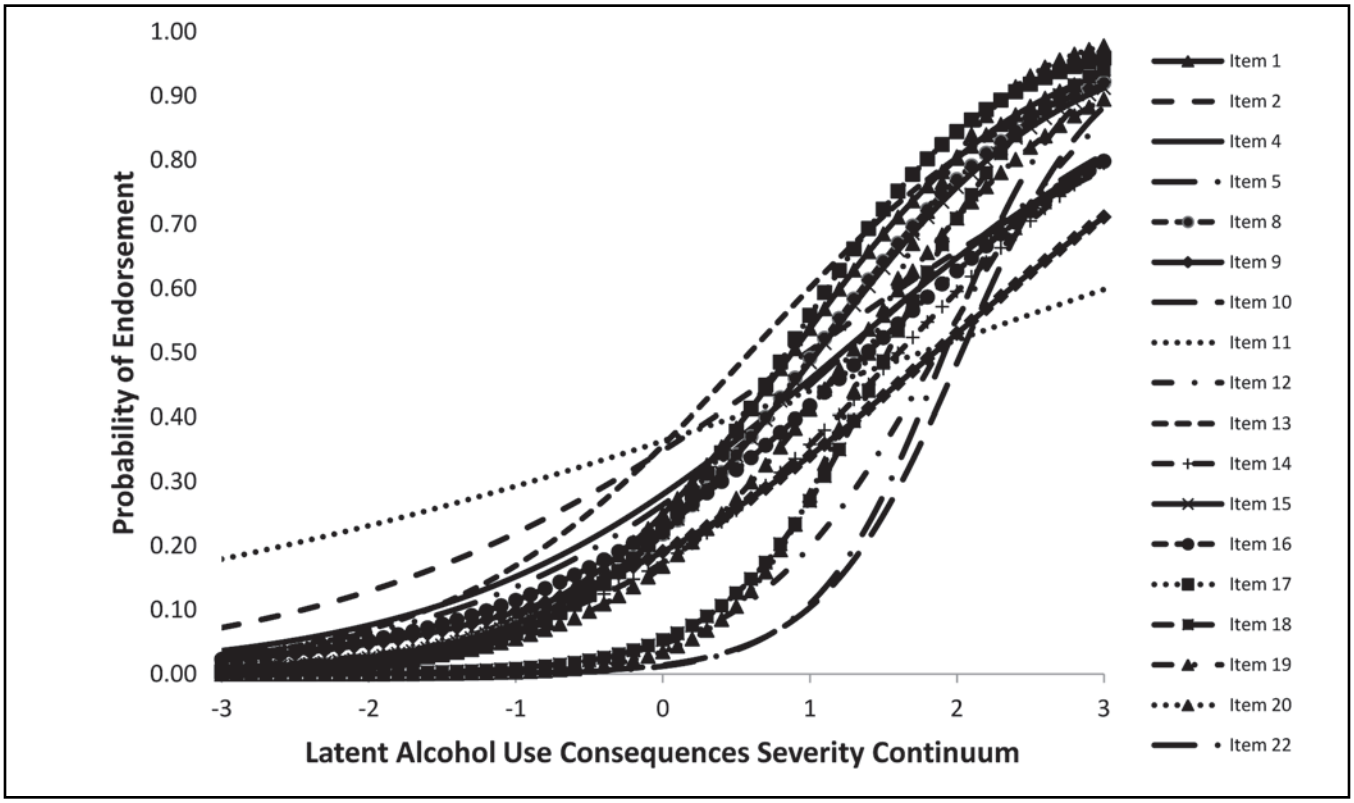

FIGURE 1. Plot of item characteristics curves for the final 18 Rutgers Alcohol Problem Index items

Table 1. Final item response theory analysis on the modified "best fitting" 18-item Rutgers Alcohol Problem Index

Item	(*n*)	% Endorsement Difficulty	*SE*	Discrim.	*SE*
10) Had withdrawal symptoms, that is, felt sick because you stopped or cut down on drinking	4.4 (14)	2.031	0.311	2.09	0.538
22) Felt psychologically or physically dependent on alcohol	6.1 (19)	1.915	0.145	2.291	2.391
12) Felt that you had a problem with alcohol	7.9 (25)	1.869	0.16	1.604	0.631
9) Tried to control your drinking by drinking at certain times of the day or certain places	14.6 (46)	1.847	0.189	0.782	0.285
11) Noticed a change in your personality	16.8 (53)	1.749	0.15	0.321	0.275
14) Tried to cut down or quit drinking	18.4 (58)	1.601	0.185	0.971	0.211
18) Had a fight, argument, or bad feelings with a relative	11.4 (36)	1.528	0.254	1.884	0.388
20) Felt you were going crazy	9.8 (31)	1.402	0.232	2.371	0.461
16) Passed out or fainted suddenly	22.9 (72)	1.391	0.145	0.854	0.226
19) Kept drinking when you promised yourself not to	18.4 (58)	1.283	0.235	1.247	0.215
5) Caused shame or embarrassment to someone	26.3 (83)	1.246	0.19	0.824	0.169
4) Went to school high or drunk	25.7 (81)	1.215	0.136	0.776	0.258
15) Suddenly found yourself in a place you could not remember getting to	27.9 (88)	1.051	0.12	1.202	0.271
8) Felt that you needed more alcohol to get the same effect	26.7 (84)	1.028	0.13	1.242	0.219
2) Got into fights, acted bad, or did mean things	30.8 (97)	0.98	0.12	0.641	0.241
1) Not able to do your homework or study for a test	32.4 (102)	0.88	0.11	1.253	0.298
17) Had a fight, argument, or bad feelings with a friend	36.8 (116)	0.841	0.128	1.458	0.221
13) Missed a day (or part of a day) of school or work	36.8 (116)	0.591	0.145	0.999	0.159

Notes: All items are sorted in descending sequence based on the difficulty parameter. Discrim. = discrimination.

"best fitting" RAPI). This secondary analysis indicated that each item provided good model fit (Figure 1).

Final item response theory model

A second two-parameter IRT analysis was conducted with the following items dropped: 3, 6, 7, 21, and 23 (Table 1). The mean number of consequences for the modified 18-item RAPI was 3.74 (*SD* = 4.18). There was a moderate degree of endorsement for each item (4.4% to 36.8%) with an overall mean of 1.36 (*SD* = 0.42) for the difficulty parameters and an overall mean of 1.27 (*SD* = 0.58) for the discrimination parameters. The values of the discrimination parameters ranged from 0.321 to 2.371 with Items 10 ("Had withdrawal symptoms, that is, felt sick because you stopped or cut down on drinking"), 20 ("Felt you were going crazy"), and 22 ("Felt psychologically or physically dependent on alcohol") having the highest discriminatory ability across the latent-trait continuum. The following items had the lowest discrimination parameters: Items 11 ("Noticed a change in your personality"), 2 ("Got into fights, acted bad, or did mean things"), and 4 ("Went to school high or drunk").

Difficulty parameter values ranged from 0.591 to 2.031 with Items 10 ("Had withdrawal symptoms, that is, felt sick because you stopped or cut down on drinking"), 22 ("Felt psychologically or physically dependent on alcohol"), and 12 ("Felt that you had a problem with alcohol") having the highest values. This implies that greater levels of negative consequences of use are necessary to respond to these items. The following items had the lowest difficulty parameter values—Items 1 ("Not able to do your homework or study for a test"), 13 ("Missed a day or part of a day of school or work"), and 17 ("Had a fight, argument, or bad feelings with a friend")—indicating that lower levels of the latent trait of alcohol-related consequences are necessary to respond to these items. Lower difficulty parameters indicate that these items were more likely to be endorsed across the sample and that lower levels of the latent trait were necessary to respond to these items.

Total information curves

The total information curves for the 23-item and 18-item RAPI appear to peak at the same location (i.e., higher end) of the continuum (Figure 2); thus, dropping five items did not appear to result in a loss of psychometric information.

Classical test theory analyses

Additional CTT analyses focused on examining the reliability and validity of the original 23-item RAPI (*M* = 30.86; *SD* = 10.73) and comparing these statistics with estimates of the modified 18-item RAPI (*M* = 23.93; *SD* = 8.36). Analyses showed that the original RAPI demonstrated a high degree of internal consistency (α = .921). Cronbach's α with each item removed (Table 2) was calculated for each of the items on the original RAPI, all of which remained in the high range with little variation (range: .912–.919). Item-to-total scale correlations on the original RAPI showed correlations ranging from *r* = .309 (Item 7, "Relatives avoided you") to *r* = .686 (Item 1, "Not able to do your homework or study for a test"). With respect to the modified 18-item

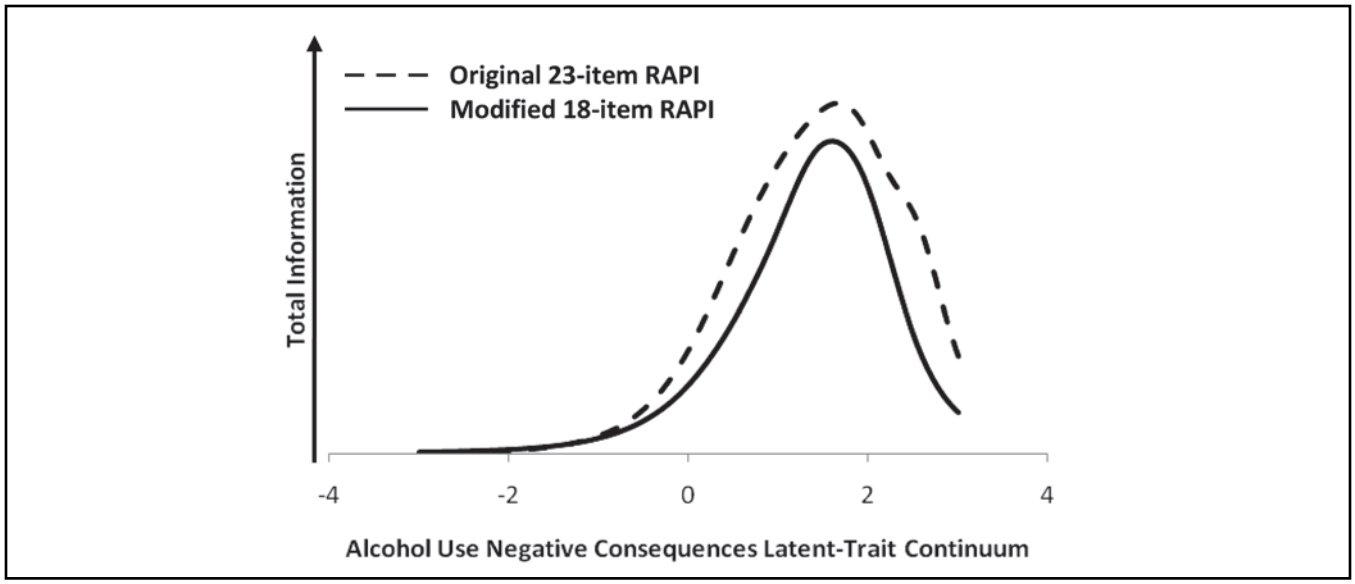

FIGURE 2. Plot of total information curves between original 23-item Rutgers Alcohol Problem Index (RAPI) and the modified 18-item RAPI. The vertical axis represents total test information. The horizontal axis represents level of ability that is measured by the underlying construct (i.e., alcohol-related negative consequences severity).

RAPI, Cronbach's coefficient α with each item removed and item-to-total scale correlations revealed negligible differences when compared with the reliability estimates obtained from the original 23-item RAPI.

Correlation coefficients indicated that the total score on the 23-item RAPI was significantly related to frequency of alcohol use (r = .56), quantity of alcohol use (r = .54), heaviest amount of alcohol consumed (r = .53), average number of drinks consumed on a typical weekday (r = .51), and average number of drinks consumed on a typical weekend (r = .47) (all ps < .001). The principal components analysis showed a single dominant factor with item loadings ranging

TABLE 2. Item-to-total scale correlations and Cronbach's α if item deleted between the original 23-item RAPI and modified 18-item RAPI

	Cronbach's α if item deleted		Item-to-total scale correlations	
Item	23-item RAPI	18-item RAPI	23-item RAPI	18-item RAPI
1) Not able to do your homework or study for a test	.912	.887	.686	.674
2) Got into fights, acted bad, or did mean things	.915	.892	.557	.552
3) Missed out on other things because you drank too much alcohol	.916	[a]	.495	[a]
4) Went to school high or drunk	.913	.888	.667	.659
5) Caused shame or embarrassment to someone	.917	.894	.481	.464
6) Neglected your responsibilities	.913	[a]	.667	[a]
7) Relatives avoided you	.919	[a]	.309	[a]
8) Felt that you needed more alcohol to get the same effect	.913	.889	.649	.636
9) Tried to control drinking	.917	.894	.478	.482
10) Had withdrawal symptoms	.918	.897	.359	.345
11) Noticed a change in your personality	.916	.894	.489	.485
12) Felt that you had a problem with alcohol	.916	.893	.594	.587
13) Missed a day (or part of a day) of school or work	.914	.891	.615	.601
14) Tried to cut down or quit drinking	.916	.893	.521	.503
15) Suddenly found yourself in a place you could not remember getting to	.913	.889	.638	.635
16) Passed out or fainted suddenly	.917	.894	.462	.474
17) Had a fight, argument, or bad feelings with a friend	.914	.891	.604	.591
18) Had a fight, argument, or bad feelings with a relative	.917	.895	.463	.454
19) Kept drinking when you promised yourself not to	.914	.891	.603	.595
20) Felt you were going crazy	.916	.893	.561	.544
21) Had a bad time	.913	[a]	.667	[a]
22) Felt psychologically or physically dependent on alcohol	.915	.892	.636	.641
23) Was told by a friend or neighbor to stop or cut down drinking	.914	[a]	.614	[a]
Final Cronbach's α	.919	.897		

Notes: RAPI = Rutgers Alcohol Problem Index. [a]Items from the 23-item RAPI that were dropped in the 18-item RAPI.

Table 3. Mean scores on six measures of alcohol use as a function of number of alcohol-related consequences on the modified Rutgers Alcohol Problem Index (in quartiles)

Post hoc comparison group	No. of alcohol-related consequences								
	(1) None (n = 93; 30.5%)		(2) 1–2 (n = 59; 19.7%)		(3) 3–5 (n = 62; 21.6%)		(4) ≥6 (n = 85; 28.3%)		
Alcohol measure	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	Post hoc
Number of drinking days	3.41	5.98	11.68	11.8	18.77	13.43	32.1	19.6	4 > 3 > 2 > 1
Frequency index	4.42	1.78	6.88	2.38	8.12	2.11	9.75	2.38	4 > 3 > 2 > 1
Quantity index	4.41	1.83	6.81	1.71	8.12	2.13	9.54	2.57	4 > 3 > 2 > 1
Typical week	0.51	1.26	1.47	1.83	2.84	2.45	3.89	2.82	4 > 3 > 2 > 1
Typical weekend	1.65	2.51	4.47	2.85	5.74	2.92	8.17	6.08	4 > 3, 2 > 1
Heaviest amount	2.55	3.32	6.43	2.81	8.21	2.97	11.92	8.42	4 > 3, 2 > 1

Notes: The numbers in parentheses in column heads refer to the numbers used for illustrating significant differences in the last column titled "Post hoc." Total N = 299, and ns for each category may not add up to total sample size because of missing data on dependent variables. The following dependent variables had missing data: n = 9 for number of drinking days; n = 4 for heaviest amount; n = 3 for typical week; n = 1 for frequency index; n = 1 for quantity index; and n = 1 for typical weekend.

from .436 to .721 that accounted for approximately 37.66% of the variance.

The 18-item RAPI showed comparable construct validity; it was significantly correlated with frequency of alcohol use (r = .56), quantity of alcohol use (r = .53), heaviest amount of alcohol consumed (r = .52), average number of drinks consumed on a typical weekday (r = .51), and average number of drinks consumed on a typical weekend day (r = .46) (all ps < .001). Last, the principal components analysis on the 18-item RAPI showed a single dominant factor (i.e., factor loadings ranging from .436 to .742) accounting for 38.36% of the variance. This provides evidence that the modified RAPI captures approximately the same amount of variance as the original version.

Follow-up analyses

We conducted exploratory analyses to determine whether alcohol consequences, as measured by the modified 18-item RAPI, were indeed associated with increased consumption. We divided the sample into four approximately equivalent groups based on the number of alcohol consequences. Each quartile represented increasing severity of alcohol problems: no consequences, one to two consequences, three to five consequences, and six or more consequences. We then conducted a multivariate analysis of variance test to examine mean differences among the four groups on each of the six drinking variables: number of drinking days, alcohol use frequency, alcohol use quantity, typical weekday consumption, typical weekend consumption, and heaviest amount of alcohol consumed. Tukey's post hoc test for planned comparisons was used (Table 3).

Omnibus results were significant, Pillai's Trace = .62, $F(18, 876) = 12.65$, $p < .01$. All Univariate F tests were significant (all ps < .01). Post hoc analyses were significant among all the groups (all ps < .05), with one exception: There were no significant differences between individuals reporting one to two consequences and those reporting three to five consequences on typical weekend consumption (p = .19) and heaviest amount of alcohol consumed (p = .23). All other findings indicated that women who reported more consequences drank alcohol more frequently and in greater quantities than those with fewer to no consequences.

Discussion

The current study examined item-level functioning of the RAPI (White and Labouvie, 1989) in first-year college women using IRT methods to investigate the severity of specific types of consequences of drinking experienced by this group. IRT analyses indicated that several items did not fit the a priori specified two-parameter model(s): Items 3 ("Missed out on other things as a result of drinking"), 6 ("Neglected your responsibilities"), 7 ("Relatives avoided you"), 21 ("Had a bad time"), and 23 ("Was told by a friend or neighbor to stop or cut down on drinking"). The lack of fit with these items may reflect unique aspects of college life that are inherent to women during their first year of college. For example, women may have been less likely to be in regular contact with relatives or family members if they had moved away from home. This may explain why Item 7, reflecting difficulties with family and relatives, was one of the least likely to be endorsed in the current sample, at only 3.2%. Further, Items 3, 6, and 21 may have demonstrated item misfit because these women and their friends likely drink together on a regular basis for social and entertainment purposes; thus, consequences such as "missing out" on activities and "having a bad time" as a result of alcohol may be rare events. That college is often associated with greater freedom and decreased structure may also explain why women were less likely to report "neglecting responsibilities" as a problem from alcohol. Another explanation

for item misfit among these five consequences is that they are poor indicators of alcohol-related consequences within this population and do not increase monotonically across the latent continuum, a primary assumption that underlies IRT (Embretson and Reise, 2000).

Measurement of difficulty parameters—how likely it is that particular items will be endorsed by the sample across the latent-trait continuum of alcohol-related consequences—revealed that different consequences operated at varying degrees of severity in the sample. The lowest difficulty parameters indicated that not being able to complete homework or school assignments, missing school or classes, and having a fight with a friend as a result of drinking were more likely to be endorsed across the sample and that these items occurred at low levels of alcohol use problems. This suggests that it is common for college women in their freshman year to experience academic and interpersonal problems as a result of drinking, even at a mild degree of problem severity. This is consistent with previous IRT analyses of alcohol use consequences in college women (Neal et al., 2006). Items that provided the highest level of differentiation (discrimination parameter) described characteristics of more severe alcohol problems and were similar to items associated with alcohol dependence: feeling physically or psychologically dependent on alcohol, having withdrawal symptoms, and experiencing psychological distress as a result of alcohol use. Perhaps consequences identified by higher discrimination and difficulty parameters (e.g., experienced withdrawal symptoms, felt dependent on alcohol, going crazy, had a problem) can be useful for identifying first-year college women who are at particular risk for developing alcohol dependence and who might benefit from targeted intervention. By identifying consequences of drinking that are particularly severe—and amenable to change—we may be able to counteract the "telescoping" effect in college women who are on the trajectory toward developing an alcohol use disorder in later adulthood.

In terms of results from CTT statistics, principal component analyses revealed negligible differences comparing the amount of variance accounted for from the modified 18-item RAPI to the 23-item RAPI; and correlations with measures of alcohol consumption were the same across both versions. The internal consistency of the 18-item version was high, as were the positive associations between the 18-item RAPI with alcohol use indices. These findings, taken together, indicate that pertinent data were not lost by omitting the five poorly fitting consequences that did not appear to map onto the underlying latent-trait continuum for this subgroup. Results from the final total information curves indicated that the 18-item RAPI provided similar information as the 23-item RAPI and mapped items that were located at the higher end of the continuum, without losing psychometric information from dropping five items. Thus, the point along the underlying latent-trait continuum at which the RAPI would be deemed most reliable did not appear to be altered significantly in the modified version. To our knowledge, there is no test to determine whether the amount of information captured under the curve has been significantly altered after omitting items. Last, more items could potentially be eliminated from the RAPI because our findings indicate some redundancy, with several pairs of items having overlapping curves (Items 10 and 22; Items 4 and 5; Items 8 and 15; and Items 1 and 17).

Multivariate analysis of variance tests further indicated that alcohol problems measured by the modified 18-item RAPI increased as a function of more frequent and intense alcohol consumption. That is, women who reported a greater number of alcohol-related consequences also reported more days drinking in the prior 90 days, reported greater frequency and quantity of alcohol use overall, consumed more standard drinks on the weekends and weekdays, and consumed larger quantities of alcohol in a 24-hour period. This provides support for the construct validity of a shortened version of the RAPI and is consistent with other studies that have examined a shorter RAPI after dropping items (Earleywine et al., 2008).

It is important to discuss how our findings compare with previous IRT studies of the RAPI in college student samples. First, similar to Neal et al. (2006), Items 20 ("Felt you were going crazy") and 22 ("Felt physically or psychologically dependent on alcohol") were among the most highly discriminating items. This indicates that these items are very good at distinguishing between high- and low-severity problem-drinking women. However, contrary to prior results, Item 7 ("Relatives avoided you") showed poor fit in our model even though it had the highest difficulty parameter endorsed by women in the Neal et al. article (2006). In the present study, Item 7 had a very low rate of endorsement (3.2%), which is consistent with results from Neal et al. (2006). Thus, we take slight inconsistencies between the present study findings and those reported by Neal et al. (2006) to mean that "relatives avoided you" can indicate very high severity of problems in college women when it is endorsed at a detectable level. Whereas Neal et al. (2006) had a sample of nearly 900 college student women, ours was a smaller sample, and we may have been underpowered to detect infrequently occurring events. Further comparisons across studies indicate that, similar to the Earleywine et al. (2008) study, we found that Item 17 ("Having a fight with a friend") and problems related to academic or occupational functioning (Item 13, "Missed a day or part of a day of school or work") were endorsed at low levels of alcohol problems (low difficulty parameter). This suggests that interpersonal and academic problems commonly arise as a result of college students' drinking more generally and do not appear to tap into a high level of severity unique to women in their first year of college; most likely this is because of the highly social nature of drinking at this age. We also found that Items 4 ("Went to school high

or drunk") and 11 ("Noticed a change in your personality") had the lowest discrimination parameters, indicating that they did a poor job of discriminating between high and low levels of severity in the sample. This is somewhat consistent with Earleywine et al. (2008), who found very poor model fit for these two items and omitted them from their final model. In addition, and in line with findings reported by Earleywine and colleagues (2008), removing poorly fitting items in our model did not appreciably change the psychometric properties of the RAPI. However, it is important to note that overall comparisons between our study findings and those previously reported can only be descriptive in nature because these earlier studies used mixed-gender samples, and one examined alcohol use consequences across all years of college (Earleywine et al., 2008).

Strengths

Few studies have conducted a psychometric evaluation of the RAPI using IRT, and none have done so specifically with college women in their first year. Our results provide valuable input relevant to modeling drinking problems in this particular group. A major strength of IRT is that it does not require group comparisons to retain meaningful findings about how particular items map onto the latent continuum of a construct within a target group (Embretson and Reise, 2000). In addition, our findings are noteworthy in that they capture the assessment of alcohol-related consequences that occur during a high-risk phase of alcohol use behavior (i.e., the first year of college) and may aid in the identification of problematic drinking patterns before more severe problems develop. This has implications for future prevention efforts with women when they first enter college, such as conducting risk/needs assessments to identify those who may be in need of targeted interventions. Last, the ranking of severity for each consequence using IRT methods may be useful in providing feedback to college women within brief motivational interviewing style-based interventions.

Limitations

Several limitations should be discussed. First, findings cannot be generalized to clinical samples and to those outside of the first year of college. Concerns about the representativeness may be allayed because drinking rates reported in the current sample are similar to those from other studies of alcohol use among female college students, in general, and those in their freshman year, in particular (Hagman and Cohn, 2011; LaBrie et al., 2007; Neal and Carey, 2007; Neal et al., 2006; Snyder and Dillow, 2010). Second, we did not make comparisons in rates of drinking across semesters. However, previous studies of drinking patterns in college men and women do not show strong semester-to-semester fluctuations in consumption (Del Boca et al., 2004). Third, although including only women in this study permits a more generalizable sample for IRT, we did not compare differential item functioning (DIF) of items across gender because of the small sample of men. DIF measures item bias, or the degree to which an item is endorsed across different groups with the same ability level. An important issue that arises with DIF is that the scales will be arbitrarily different between two groups if they have different "ability" distributions (Stone and Zhang, 2003; Zumbo, 2007). In the case of alcohol use consequences, prior studies suggest that men and women may not share the same types of consequences and may have different "ability" levels (Earleywine et al., 2008; Neal et al., 2006). Thus, DIF results will be difficult to meaningfully interpret (Zumbo, 2007). Last, because we used cross-sectional data, we cannot make causal assumptions about associations between alcohol consumption and alcohol problems. It would be noteworthy to examine how and whether the underlying latent continuum of drinking-related consequences changes over the course of college and afterward.

Conclusions

Extrapolating from these data, our study findings present a "first pass" at identifying, using IRT methods, several severe problems of use that are experienced by young women during the early years of their drinking. This may shed light onto the causes and pathways of developing an alcohol use disorder and other health-risk behavior patterns in adulthood and may also provide strategies for primary prevention efforts with specific subsets of college students. Given the considerable interest in adopting gender-specific approaches to treating alcohol use behavior, knowledge of these indicators could be important for understanding what may be unique about drinking among female college students. Future analyses comparing RAPI item performance across subgroups of college-aged individuals, over time, and between college students and their non–college-attending peers would inform prevention efforts aimed at decreasing risky drinking at different developmental points.

References

Carpenter, R., Fishlock, A., Mulroy, A., Oxley, B., Russell, K., Salter, C., . . . Heffernan, C. (2008). After 'Unit 1421': An exploratory study into female students' attitudes and behaviours towards binge drinking at Leeds University. *Journal of Public Health, 30,* 8–13.

Dawson, D. A., Grant, B. F., & Harford, T. C. (1995). Variation in the association of alcohol consumption with five DSM-IV alcohol problem domains. *Alcoholism: Clinical and Experimental Research, 19,* 66–74.

Del Boca, F. K., Darkes, J., Greenbaum, P. E., & Goldman, M. S. (2004). Up close and personal: Temporal variability in the drinking of individual college students during their first year. *Journal of Consulting and Clinical Psychology, 72,* 155–164.

Earleywine, M., LaBrie, J. W., & Pedersen, E. R. (2008). A brief Rutgers

Alcohol Problem Index with less potential for bias. *Addictive Behaviors, 33,* 1249–1253.

Embretson, S. E., & Reise, S. (2000). *Item response theory for psychologists.* Mahwah, NJ: Erlbaum.

Fisher, B. S., Cullen, F. T., & Turner, M. G. (2000). *The sexual victimization of college women* (Research Report NCJ 182369). Washington, DC: U.S. Department of Justice, National Institute of Justice. Available at https://www.ncjrs.gov/pdffiles1/nij/182369.pdf

Haas, A. L., & Peters, R. H. (2000). Development of substance abuse problems among drug-involved offenders. Evidence for the telescoping effect. *Journal of Substance Abuse, 12,* 241–253.

Hagman, B. T., & Cohn, A. M. (2011). Toward DSM-V: Mapping the alcohol use disorder continuum in college students. *Drug and Alcohol Dependence.* Advance online publication. Available at http://www.sciencedirect.com/science/article/pii/S0376871611001475.

Hagman, B. T., Cohn, A. M., Noel, N. E., & Clifford, P. R. (2010). Collateral informant assessment in alcohol use research involving college students. *Journal of American College Health, 59,* 82–90.

Hagman, B. T., Kuerbis, A. N., Morgenstern, J., Bux, D. A., Parsons, J. T., & Heidinger, B. E. (2009). An Item Response Theory (IRT) analysis of the Short Inventory of Problems-Alcohol and Drugs (SIP-AD) among non-treatment seeking men-who-have-sex-with-men: Evidence for a shortened 10-item SIP-AD. *Addictive Behaviors, 34,* 948–954.

Hingson, R., & Kenkel, D. (2004). Social, health, and economic consequences of underage drinking. In R. J. Bonnie & M. E. O'Connell (Eds.), *Reducing underage drinking: A collective responsibility* (pp. 351–382). Washington, DC: The National Academies Press. Retrieved from http://books.nap.edu/openbook.php?record_id=10729&page=R1

Hingson, R. W., Heeren, T., Zakocs, R. C., Kopstein, A., & Wechsler, H. (2002). Magnitude of alcohol-related mortality and morbidity among U.S. college students ages 18-24. *Journal of Studies on Alcohol, 63,* 136–144.

Hingson, R. W., Zha, W., & Weitzman, E. R. (2009). Magnitude of and trends in alcohol-related mortality and morbidity among U.S. college students ages 18–24, 1998–2005. *Journal of Studies on Alcohol and Drugs, Supplement 16,* 12–20.

Horwitz, A. V., & White, H. R. (1987). Gender role orientations and styles of pathology among adolescents. *Journal of Health and Social Behavior, 28,* 158–170.

Johnson, P. B., Richter, L., Kleber, H. D., McLellan, A. T., & Carise, D. (2005). Telescoping of drinking-related behaviors: Gender, racial/ethnic, and age comparisons. *Substance Use & Misuse, 40,* 1139–1151.

Johnston, L. D., O'Malley, P. M., Bachman, J. G., & Schulenberg, J. E. (2004a). Monitoring the Future national survey results on drug use, 1975-2003. Volume I: Secondary school students (NIH Publication No. 04-5507). Bethesda, MD: National Institute on Drug Abuse.

Johnston, L. D., O'Malley, P. M., Bachman, J. G., & Schulenberg, J. E. (2004b). Monitoring the Future national survey results on drug use, 1975-2003. Volume II: College students and adults ages 19-45 (NIH Publication No. 04-5508). Bethesda, MD: National Institute on Drug Abuse.

LaBrie, J. W., Huchting, K., Pedersen, E. R., Hummer, J. F., Shelesky, K., & Tawalbeh, S. (2007). Female college drinking and the social learning theory: An examination of the developmental transition period from high school to college. *Journal of College Student Development, 48,* 344–356.

MacNeela, P., & Bredin, O. (2011). Keeping your balance: Freedom and regulation in female university students' drinking practices. *Journal of Health Psychology, 16,* 284–293.

Mann, K., Ackermann, K., Croissant, B., Mundle, G., Nakovics, H., & Diehl, A. (2005). Neuroimaging of gender differences in alcohol dependence: Are women more vulnerable? *Alcoholism: Clinical and Experimental Research, 29,* 896–901.

Martens, M. P., Neighbors, C., Dams-O'Connor, K., Lee, C. M., & Larimer, M. E. (2007). The factor structure of a dichotomously scored Rutgers Alcohol Problem Index. *Journal of Studies on Alcohol and Drugs, 68,* 597–606.

McCabe, S. E. (2002). Gender differences in collegiate risk factors for heavy episodic drinking. *Journal of Studies on Alcohol, 63,* 49–56.

Mohler-Kuo, M., Dowdall, G. W., Koss, M. P., & Wechsler, H. (2004). Correlates of rape while intoxicated in a national sample of college women. *Journal of Studies on Alcohol, 65,* 37–45.

Muraki, E., & Bock, R. (2002). PARSCALE (Version 4.1) [Computer software]. Lincolnwood, IL: Scientific Software International.

Neal, D. J., & Carey, K. B. (2007). Association between alcohol intoxication and alcohol-related problems: An event-level analysis. *Psychology of Addictive Behaviors, 21,* 194–204.

Neal, D. J., Corbin, W. R., & Fromme, K. (2006). Measurement of alcohol-related consequences among high school and college students: Application of item response models to the Rutgers Alcohol Problem Index. *Psychological Assessment, 18,* 402–414.

Neal, D. J., & Fromme, K. (2007). Event-level covariation of alcohol intoxication and behavioral risks during the first year of college. *Journal of Consulting and Clinical Psychology, 75,* 294–306.

Nolen-Hoeksema, S. (2004). Gender differences in risk factors and consequences for alcohol use and problems. *Clinical Psychology Review, 24,* 981–1010.

Parks, K. A., Romosz, A. M., Bradizza, C. M., & Hsieh, Y.-P. (2008). A dangerous transition: Women's drinking and related victimization from high school to the first year at college. *Journal of Studies on Alcohol and Drugs, 69,* 65–74.

Perkins, H. W. (2002). Surveying the damage: A review of research on consequences of alcohol misuse in college populations. *Journal of Studies on Alcohol, Supplement 14,* 91–100.

Piazza, N. J., Vrbka, J. L., & Yeager, R. D. (1989). Telescoping of alcoholism in women alcoholics. *International Journal of the Addictions, 24,* 19–28.

Snyder, T. D., and Dillow, S. A. (2010). *Digest of education statistics 2009* (NCES 2010-013). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Stone, C. A., & Zhang, B. (2003). Assessing goodness of fit of item response theory models: A comparison of traditional and alternative procedures. *Journal of Educational Measurement, 40,* 331–352.

Sugarman, D. E., DeMartini, K. S., & Carey, K. B. (2009). Are women at greater risk? An examination of alcohol-related consequences and gender. *The American Journal on Addictions, 18,* 194–197.

Testa, M., Hoffman, J. H., & Livingston, J. A. (2010). Alcohol and sexual risk behaviors as mediators of the sexual victimization-revictimization relationship. *Journal of Consulting and Clinical Psychology, 78,* 249–259.

Ullman, S. E. (2007). Mental health services seeking in sexual assault victims. *Women & Therapy, 30,* 61–84.

U.S. Department of Health and Human Services. (2007). *The Surgeon General's Call to Action to Prevent and Reduce Underage Drinking.* Rockville, MD: Author. Available at http://www.surgeongeneral.gov/topics/underagedrinking/calltoaction.pdf

Wechsler, H., Lee, J. E., Kuo, M., Seibring, M., Nelson, T. F., & Lee, H. (2002). Trends in college binge drinking during a period of increased prevention efforts. Findings from 4 Harvard School of Public Health College Alcohol Study surveys: 1993–2001. *Journal of American College Health, 50,* 203–217.

White, H. R., & Labouvie, E. W. (1989). Towards the assessment of adolescent problem drinking. *Journal of Studies on Alcohol, 50,* 30–37.

Zumbo, B. D. (2007). Three generations of DIF analyses: Considering where it has been, where it is now, and where it is going. *Language Assessment Quarterly, 4,* 223–233.